PHD

Modular Normalisation of Classical Proofs

Ralph, Benjamin

Award date:
2019

Awarding institution:
University of Bath

Link to publication

Modular Normalisation of Classical Proofs

submitted by

BENJAMIN RALPH

for the degree of

DOCTOR OF PHILOSOPHY

of the

UNIVERSITY OF BATH

DEPARTMENT OF COMPUTER SCIENCE

February 2019

Signature of Author ....

Benjamin Ralph

Abstract

The main contribution of this thesis is to present a study of two normalisation theorems and proofs in classical logic: one propositional, one first-order. For propositional logic, we show a local cycle removal procedure through reductions on merge contractions that ensures that proofs can be decomposed—that contractions can be pushed to the bottom of a proof—in a straightforward way. For first-order logic, we show how decomposition of proofs can correspond to two presentations of Herbrand's Theorem, and how we can use translations into expansion proofs to give a new, indirect cut elimination theorem for first-order logic.

In addition, an old but interesting cut elimination method for propositional logic, the experiments method, is formally presented for the first time, and we extend the theory of merge contractions to first-order logic.

Acknowledgements

Over the course of my PhD, the supervision of Alessio Guglielmi—a sort of holy fool, for science—has been a strange kind of delight: at times exasperating, at times enlightening, but never, ever dull. Most importantly, Alessio gave me what must have been the correct balance of advice and space to enable me to grow as a mathematician and a person, for which I will be forever grateful. To borrow his own words about his own supervisor: scientists and teachers like Alessio are very rare and essential. I feel lucky to have met him.

I feel privileged that Andrea Aler Tubella has turned from a mentor to a collaborator, while also becoming a good friend. Working with her has never felt like a chore, and has made me a much better mathematician.

Anupam Das has been an inspiration and a friend, always generous and enthusiastic. At the beginning of my PhD, a daunting and difficult time, his advice was comforting and invaluable.

Visits to Michel Parigot in Paris with Alessio and Andrea to work on the first-order calculus with indicated substitutions were also hugely formative and instructive. It's a regret that only a hint of that work has made it into this thesis.

The wider deep inference community—Paola, Fanny, David, Alessio, Michel, Tom, Lutz, Kai, Sonia, and others—have been a welcoming family both in Bath and further afield. Likewise, the Mathfound group at Bath and the wider Computer Science department have provided a stimulating and collegiate environment in which to work. My two examiners, Willem Heijltjes and Georg Moser were a model of academic rigour and enthusiasm, engaging critically with my work to help bring out the best thesis possible.

I'd like to thank my two flatmates, Sam and Kat, for indulging my quirks and habits, and being such good company. I'll miss my friends in Bath too, who made me feel at home in a new city. I'll also miss all those I met through Bath Labour and Momentum, the local UCU branch, and Bath Students Against Fees and Cuts, who, against the odds, made the city and university feel like living, breathing, caring communities. Thanks also to Raisa, Luis and Lupe for making New York feel like home whenever I visited.

I must also thank Jack for influencing this thesis more than he could have possibly imagined by introducing me to Herbrand's Theorem, and patiently listening to my garbled explanations of my work on long walks.

I lived with my grandparents, David and Manon, for the final months in which this thesis was written, and it's remarkable how fondly I'll look back on this short period, given how stressful it should have been. Finally, I would like to thank my parents, Chris and Anna, who have been a rock of love and support.

Contents

List of Figures

Introduction

> 'Hence there can never be surprises in logic.'
>
> L. Wittgenstein [Wit61]

The fundamental compression mechanism for sequent calculus proofs is understood to be the cut; therefore the proof theoretic analogue to computation is seen to be cut elimination. However, in classical logic—both propositional and first-order—the complexity of cut elimination is greatly determined by the position and behaviour of contractions in proofs. In particular, the centrality of contractions on existential formulae to the structure of first-order proofs has been highlighted in the body of the work that was initiated by Jacques Herbrand in 1930 [Her30; Her71]. The fundamental result shows that all true first-order formulae can be proved in two phases: a primary stage, in which only certain contractions are allowed; followed by the rest of the proof. Crucially, a bound on the contractive phase of the proof renders finding the rest of the proof a decidable problem: contraction, therefore, can be seen as the locus of undecidability in classical first-order logic.

Why then the traditional focus on cuts over contractions in traditional proof theory? A compelling answer looks to the 'Two Restrictions on Contraction' identified by Kai Brünnler for sequent systems with multiplicative rules: that contractions cannot be reduced to atomic form and that the contractive phase of the proof cannot be isolated from the rest [Brü03b]. Move to a deep inference proof system, however, and these restrictions are lifted. Thus, whole new avenues for proof theoretic inquiry are unlocked by a move to deep inference, with classes of proofs and normal forms that were previously inexpressible, and a finer level of analysis of processes of normalisation made possible. In particular, we have classes of proofs, such as those guaranteed by Herbrand's Theorem, which are decomposed into contractive and non-contractive phases. In deep inference systems for classical logic, every provable formula has a decomposed proof; there are procedures that allow us to decompose all proofs into this form. But what is a deep inference proof system?

A Brief History of Deep Inference

Deep inference started off as a personal research project of Alessio Guglielmi, studying an 'extreme form of linear logic' that extends the notion of linearity beyond the formula-level to the level of the proof system itself [Gug]. Syntactically, this is achieved by allowing the same means of composition for derivations as formulae—obliterating the 'object-level/meta-level', or 'structural connective/logical connective' distinction common to proofs systems in the Gentzen tradition. The theoretical direction of travel can be thought of as opposed to that which begat hypersequent and display calculi [Avr87; Bel82; CRW14]; deep inference is characterised by a reduction not a proliferation of (structural) syntax. However an immediate motivation was the same: just as extensions of the sequent calculus allow for more logics to be captured by cut-free systems, it was very early on discovered that deep inference could be used to capture a linear logic extended with process-algebra's sequential composition, BV, conjectured to be equivalent to Christian Retoré's Pomset Logic and provably inexpressible in the sequent calculus [Gug07; Ret97; Tiu01; Tiu05].

The initial phase of development took place in Dresden between 1999 and 2003, culminating in one Master's and two PhD theses that established the subject [Brü03a; Str03; Tiu01]. This first wave of research in deep inference proof theory is perhaps epitomised by the discovery of proof systems with fully local structural rules for classical [Brü06a; BT01], intuitionistic [Tiu06], linear [Gug07; Str03], and modal [Brü09; SS04] logics. The ability to reduce cut to atomic form led to whole new methods of cut elimination, such as splitting [Brü03a; Gug07], while the reduction of contraction to atomic form was key to work on computational interpretations of deep inference, such as the atomic lambda calculus, a typed calculus with explicit sharing [GHP13]; the atomic lambda-mu calculus, [Par92; He18]; and a Curry-Howard correspondence between interaction nets and a deep inference system for MELL [GM13].

The next wave of research in deep inference, at least for classical propositional logic, was initiated by the discovery of the atomic flow by Tom Gundersen. The idea is similar to that of Samuel Buss's *logical flow graphs* [Bus91; Car02], which were developed to reason about sequent calculus proofs ignoring certain inessential aspects of the proof, what Jean-Yves Girard described as the 'bureaucracy of syntax' [Gir89]. An atomic flow is a directed acyclic graph with each node corresponding to an atomic structural rule and vertices corresponding to atoms. Composition of flows matches composition of open deduction derivations: we can adjoin two flows vertically (assuming the respective inputs and outputs match) or horizontally. However, unlike proof nets [Gir87; LS05], atomic flows are not proof systems—there is no polynomial-time correctness condition for deciding whether a certain flow guarantees the proof of a proposition.

The first major result for atomic flows was showing that normalisation of propositional proofs can be carried out at the level of the flow, with rewrites of flows descending cleanly to rewrites of derivations [GG08; Gun09]. Global flow transformations, in particular the *path breaker*, ensured that the rewriting systems for the flows terminated [GGS10], with this approach being refined by the use of threshold formulae to show that cut elimination need only increase the size of a proof quasi-polynomially, with polynomial size cost conjectured [Bru+16; Jeř08]. More recently, further work on the proof complexity of propositional logic has been carried out by Anupam Das and others, with at least four interesting results in the following papers [Das14b; Das14a; Das15].

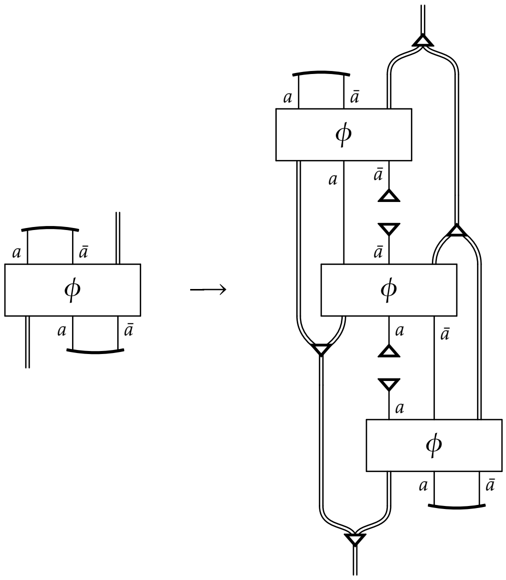


Figure 1: The 'path breaker' for a derivation with one atom. For technical details see [GGP10].

In particular, the crucial role of contraction (and cocontraction) in managing the complexity of propositional proofs has been explicated, with contraction-cocontraction pairs, known as 'sausages', key to proof compression [Das15].

Most recently, the development of subatomic logic, chiefly by Andrea Aler Tubella and Guglielmi, has unified the proof theory of a wide variety of logics under a single scheme [AT17; ATG18]. By considering atoms not as fundamental objects of study for propositional proofs, but as self-dual non-commutative connectives that can be interpreted back into a standard system, one can fit structural, logical and equality rules for classical, linear, non-commutative and modal logics into a remarkably uniform rule-shape. Furthermore, the regularity of the rule-shape enables normalisation procedures—decomposition and splitting—that can be described independent of a particular logic and applicable to many at once.

Decomposition via Cycle Removal

The first part of this thesis looks at how the techniques developed for subatomic systems can be specialised to classical propositional logic. The key obstacle to normalization is the existence of infinite reduction paths for rewriting systems that decompose proofs. Above, we described how global transformations on atomic flows can be used to convert derivations into forms where local rewriting of flows is terminating. They do so by ensuring that there are no 'cycles' between identities and cuts in flows, which, if they contain contractions, can lead to infinite reduction sequences. Thus, we obtain the separation of cut

$$\mathsf{m}\,\frac{(A\{a\} \wedge B) \vee (C \wedge D\{\bar{a}\})}{(A\{a\} \vee C) \wedge (B \vee D\{\bar{a}\})}$$

Figure 2: A 'critical medial' rule for a cycle

elimination into three separate procedures:

1. Cycle elimination
2. Decomposition of proofs into contractive and non-contractive phases.
3. Elimination of cuts in the linear fragment via splitting.

Unlike decomposition and splitting, cycle elimination via global transformations involves plugging parts of derivations into standard moulds, such as the 'path breaker' [GGP10]. These moulds provide a clean way to guarantee termination without the need to 'look inside' the derivation. In contrast, the reductions for decomposition involve permuting individual atomic contractions and cocontractions around the derivation. The question then raised is whether we can remove cycles in a similar way to decomposition, as a fully local procedure carried out on atomic inference rules. Not quite: although the procedure we show is 'internal' rather than 'external', it is not fully atomic. The procedure we show stems from the observation that a certain linear rule, the 'medial' rule, is the only possible inference rule that is able to change the connective linking an atom and its dual from disjunction to conjunction, a necessary ingredient for a cycle. Thus, if we can permute this 'critical' medial rule to the bottom of the cycle, we can 'break' the cycle against the cut. Unfortunately, it is not as straightforward to permute medial rules through a derivation as it is for certain other rules, in particular structural rules.

Merge Contractions

To overcome the problem of permuting the medial rule through proofs, a new rule generalising medials and atomic contractions is defined, the 'merge contraction', which can be permuted down to break the cycle. Once some theory for merge contractions is developed, the cycle removal procedure itself becomes simple. What is a merge contraction, though? One way to think about the underlying idea is by (coinductively) asking the question 'to what extent can two formulae A and B be contracted?'. If at least one of the two formulae is atomic, the question can be answered immediately: 'completely' or 'not at all'. However if our two formulae are non-atomic, and have the same top-level connective, we can reduce the question to 'to what extent can the maximal subformulae of A and B be contracted?'. Ignoring certain details that will be covered in the body of the thesis, this is the basic idea behind the definition of

$$\mathsf{mc}{\downarrow}\frac{a \vee a}{a} \quad \mathsf{mc}{\downarrow}\frac{A \vee \mathsf{f}}{A} \quad \mathsf{mc}{\downarrow}\frac{(A \wedge B) \vee (A \wedge C)}{A \wedge (B \vee C)}$$

$$\mathsf{mc}{\downarrow}\frac{(A \wedge B) \vee (\mathsf{f} \wedge (C \wedge D))}{A \vee (B \vee (C \wedge D))} \quad \mathsf{mc}{\downarrow}\frac{(\exists x A \wedge \forall y B) \vee (\exists x A \wedge \forall y D)}{\exists x A \wedge \forall y (B \wedge D)}$$

Figure 3: Examples of merge contractions

the 'merge set' of A and B: the set of formulae that can be obtained by contracting/merging A and B, including at a minimum $A \vee B$. By allowing any element of this merge set, C to be the conclusion of the merge contraction inference rule $\mathsf{mc}{\downarrow}\frac{A \vee B}{C}$, it turns out we can encapsulate an enormous amount of 'contractive' behaviour in a single inference rule: atomic contractions, medials, distributivity, unit equations, as well as a whole class of structured derivations combining these inference rules. Furthermore, it turns out that the rewriting system for atomic contractions that allows us to decompose proofs can be generalised to merge contractions, for both propositional and first-order classical logic.

With this rewriting system for merge contractions, we are able to provide a new proof for cycle elimination, similar but different to those in [ATGR17; AT17]. However, it is expected that this new proof is merely the first application of the concept, especially since we show that the rewriting system used can be straightforwardly extended to first-order logic. A categorical treatment of merge contractions is an obvious next step, as well as more ambitious normalisation theorems for the merge contraction rewriting system than those presented in this thesis.

Herbrand's Theorem

In the second part of this thesis, we look at decomposition through the lens of Herbrand's Theorem. Like much of the development of first-order proof theory, the work of Herbrand was done in the context of Hilbert's Program, an attempt by (overlapping) sections of the mathematical and philosophical communities to rebuild faith in set theory, to retake 'the paradise that Cantor created for us' by formalising and proving the consistency of infinitary mathematics by finitary means [Hil25]. Since a proof of consistency would be provided by a sound, complete and decidable proof system for first-order arithmetic, many leading mathematicians, including Herbrand, focussed their energies on the newly-emerged field of proof theory. In essence, Herbrand's project was to find a concise representation of the content of first-order proofs that was truly first-order, as opposed to merely propositional; to isolate the minimal amount of non-propositional information that is contained in a first-order proof. The importance of such a project, especially to the then embryonic field of theoretical computer science, is that while propositional logic is decidable, full first-order logic is not. With this in mind, Hebrand's Theorem can be

$$\begin{array}{ccc}
 & \exists a,b \in \mathbb{R}(\overline{\mathbb{Q}}(a) \wedge \overline{\mathbb{Q}}(b) \wedge \mathbb{Q}(a^b)) & \\
\swarrow & & \searrow \\
\exists a,b(\overline{\mathbb{Q}}(a) \wedge \overline{\mathbb{Q}}(b) \wedge \mathbb{Q}(a^b)) & \vee & \exists a,b(\overline{\mathbb{Q}}(a) \wedge \overline{\mathbb{Q}}(b) \wedge \mathbb{Q}(a^b)) \\
\downarrow & & \downarrow \\
\overline{\mathbb{Q}}(\sqrt{2}) \wedge \overline{\mathbb{Q}}(\sqrt{2}) \wedge \mathbb{Q}(\sqrt{2}^{\sqrt{2}}) & \vee & \overline{\mathbb{Q}}(\sqrt{2}^{\sqrt{2}}) \wedge \overline{\mathbb{Q}}(\sqrt{2}) \wedge \mathbb{Q}(2) \\
\downarrow & & \downarrow \\
\mathbb{Q}(\sqrt{2}^{\sqrt{2}}) & \vee & \overline{\mathbb{Q}}(\sqrt{2}^{\sqrt{2}})
\end{array}$$

Figure 4: A demonstration of $\exists a,b \in \mathbb{R}(\overline{\mathbb{Q}}(a) \wedge \overline{\mathbb{Q}}(b) \wedge \mathbb{Q}(a^b))$ in 'Herbrand' style.

seen as teasing out the kernel of undecidability from first-order logic.

To illustrate Herbrand's Theorem, simple examples can be elucidatory. Take the first-order translation of the sentence 'There exists two irrational numbers a and b such that a^b is rational.', which we will abbreviate $(\overline{\mathbb{Q}}(a) \wedge \overline{\mathbb{Q}}(b) \wedge \mathbb{Q}(a^b))$ For an intuitionist, a proof of the statement would have to be a pair of rationals (a,b) that fit the bill. However, classical proofs can be more liberal: if for some finite list of pairs $(a_1,b_1),\ldots,(a_n,b_n)$, we can prove that $\bigvee_1^n(\overline{\mathbb{Q}}(a_i) \wedge \overline{\mathbb{Q}}(b_i) \wedge \mathbb{Q}(a_i^{b_i}))$ is a true sentence, then we have proved the original statment. It turns out that the second approach gives us a simpler proof than the first. For if we choose $a_1 = \sqrt{2}$, $b_1 = \sqrt{2}$, $a_2 = \sqrt{2}^{\sqrt{2}}$, $b_2 = \sqrt{2}$, proving the required statement is simply shown to reduce to proving the classical propositional tautology that $\sqrt{2}^{\sqrt{2}}$ is either rational or irrational. This strategy of proving existential formula by expanding them into a finite disjunction of instantiations is the crux of Herbrand's Theorem. Unfortunately, Herbrand's own statement of his theorem, let alone the proof, is notoriously 'hard to follow' [HB74; Web14], with most modern treatments reformulating the material. For example, we have the following statement of Herbrand's theorem in a prominent, more recent exposition by Buss [Bus95]:

Herbrand's Theorem. *A first-order formula A is valid if and only if A has a Herbrand proof. A Herbrand proof of A consists of a prenexification $A^\star$ of a strong $\vee$-expansion of A plus a witnessing substitution σ for $A^\star$.*

Herbrand's theorem has also been stated and proven in a deep inference setting, in [Brü06a]:

Herbrand's Theorem. *For each proof of a formula S in system* SKSgr *there is a substitution σ, a propositional formula P, a context $Q\{\ \}$ consisting only of quanti-*

fiers and a Herbrand proof*:*

$$
\begin{array}{c}
\left.\begin{array}{c}-\end{array}\right\Vert \mathsf{KS} \cup \{\mathsf{ai}\uparrow\} \\
\forall \vec{x} P \sigma \\
\Vert \{\mathsf{n}\downarrow\} \\
Q\{P\} \\
\Vert \{\mathsf{gr}\downarrow\} \\
S' \\
\Vert \{\mathsf{qc}\downarrow\} \\
S
\end{array}
$$

From these we can abstract a pattern of four key steps necessary for a Herbrand Proof.

1. Expansion of existential subformulae.
2. Prenexification/elimination of universal quantifiers.
3. Term assignment.
4. Propositional tautology check.

This strategy is common to the two approaches. But we can also note the difference between the two formulations. One key difference between the two is that, while Buss's definition of a Herbrand proof is that it is a *sui generis* form of proof, not a particular class of proof in a particular proof formalism, whereas Brünnler's is merely a subclass of proofs in a particular deep inference system. In deep inference, each of the four conditions for the Herbrand Proof correspond to certain first-order inference rules, rather than an *ad hoc* operation on a first-order formula. Is this not possible in the sequent calculus?

Decomposition as Herbrand's Theorem

The key to the difference between Buss's and Brünnler's Herbrand proofs is due to one of the properties of proofs that Brünnler showed are possible in deep inference but not sequent calculus systems. The second property is the following:

'Proofs can be separated into two phases (seen bottom-up): The lower phase only contains instances of contraction. The upper phase contains instances of the other rules, but no contraction. No formulae are duplicated in the upper phase.' [Brü03b]

Brünnler shows that a standard sequent calculus proof system with multiplicative rules cannot satisfy this property. The suggested way round this restriction is to use systems with *deep contraction*. In fact, this restriction on sequent calculus systems is shown by Richard McKinley in [McK10] to create a gap in Buss's proof of Herbrand's theorem in [Bus95]. The faulty proof assumes that if one restricts contraction to only existential formulae, one retains completeness (assuming a multiplicative $\wedge R$ rule). That this is false can be seen by considering the sequent below, where the application of any multiplicative $\wedge R$ rule leads to an invalid sequent:

$$\vdash \forall x A \wedge \forall x B, \left(\exists x \bar{A} \vee \exists x \bar{B}\right) \wedge \left(\exists x \bar{A} \vee \exists x \bar{B}\right)$$

It is the inability of sequent systems to satisfy this proper that ensures that Herbrand proofs can never be expressed as a subclass of sequent proofs. Moreover, the first stage of a Herbrand proof is duplicating existential formulae, which when translated into a bottom-up proof system is performed by contraction. Therefore Herbrand proofs, in common with decomposed proofs, have contractions at the bottom of their proofs; we can see Herbrand's theorem as the first-order instantiation of the more general proof theoretic procedure of decomposition.

Another advantage that the move to deep inference gives us is the possibility to dramatically reduce the size of 'Herbrand Disjunction', the midsequent between the first-order and propositional part of the proof. Using the lower bounds on Herbrand expansion discovered by Richard Statman [Sta79], Matthias Baaz and Alexander Leitsch showed that there are certain first-order formulae whose prenexification increases the size of their Herbrand Expansion nonelementarily [BL94]. Conversely, there are certain formulae for which deprenexification reduces the size of their Herbrand Expansion nonelementarily. However, deprenexification is clearly not possible for proof systems with shallow inference, since quantifiers are pushed deep inside formulae. Baaz and Juan Aguilera have shown that one can recover the complexity advantage by a sequent system with locally unsound inference rules [AB16]. Yet, as they note, including a family of deep inference rules that they call 'quantifier shifts' (but were already present in Herbrand's work as the 'rules of passage') is sufficient for the speed-up, with no need for any local unsoundness. Thus, by adding these rules of passage to a cut-free deep inference system, we obtain the nonelementary proof-size reduction for certain formulae.

Decomposed Proofs as Expansion Proofs

Herbrand proofs are not the only way that Herbrand's theorem has been reinterpreted. Another strand of research was initiated by the definition of 'expansion proofs', which involve a generalisation of Herbrand's Theorem to higher-order logics [Mil87]. The idea is to enrich formulae, explicitly adding in contraction and instantiation information as syntax, so that they contain the Herbrand disjunction attached to each first-order formulae as their 'deep formula'.

One intuitive way to think about expansion proofs is as representations of Coquand-style games [Coq95]: ∃loise, who can choose terms at existential node, plays ∀belard, who chooses variables at universal nodes. Once all the quantifiers are expanded, ∃loise wins if the resulting propositional formula is a tautology, ∀belard otherwise. The first-order formula is true iff ∃loise has a winning strategy.

However, the game described above is not complete for classical logic. For classical proofs, we must give ∃loise the ability to 'backtrack', returning to any previously expanded existential node at any point to choose another term. The winning condition is now a disjunction over all of ∃loise's choices. Since ∃loise can only include free variables in her terms once ∀belard has played them, this gives her access to more winning strategies, matching the fact that more first-order sentences are true classically than intuitionistically.

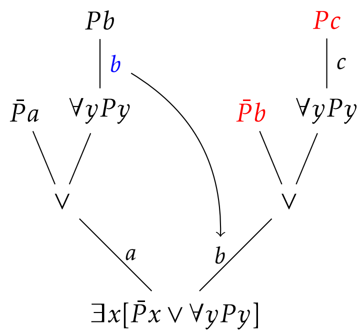


Figure 5: An expansion proof of the drinker's formula

As an example, we consider the drinker's formula, $\exists x \forall y[\bar{P}x \vee Py]$, as popularised by Smullyan: 'There is someone in the pub such that, if they are drinking, then everyone in the pub is drinking.' As the outermost quantifier is an existential, ∃loise moves first. At this point, there is no other move but to choose a closed term at random. ∀belard then chooses a variable to play—clearly he should not pick the same term as ∃loise. Assuming not, the resulting tautology would at this point be something along the lines of $Pa \vee \bar{P}b$, it seems as if ∃loise does not have a winning strategy. However, ∃loise is allowed to backtrack, choosing the variable ∀belard picks for her second existential witness. This time, whatever ∀belard picks, the disjunction over the two choices will be a tautology, say $\left(Pa \vee \bar{P}b\right) \vee \left(Pb \vee \bar{P}c\right)$.

In the original presentation of expansion proofs, Miller provides translations back and forth between his new formalism and the sequent calculus. However expansion proofs did not enjoy all the usual features of a proof system. Firstly, there is no account of the propositional aspect, just a tautology check. Obviously one could be given, but there is no natural analogue of expansion proofs for classical propositional logic. This isn't really a problem—the motivation behind expansion proofs is certification of first-order proofs, and using a first-order proof as a certificate for itself isn't of much use. Secondly, there is no means to compose proofs by cut and certainly no cut elimination.

In fact, proving cut elimination for expansion proofs (or similar structures) has been a relatively active topic of research in recent years. In [Hei10], a system of 'proof forests' is presented, a graphical formalism of expansion proofs with cut and a local rewrite relation that performs cut elimination. Similar work has been carried out in [McK13] and more recently in [AHW18]. A more categorical approach has also been given in [Alc+17]. Instead of a *sui generis* formalism for expansion proofs with cut, we show a class of deep-inference proofs that closely correspond to expansion proofs and another to expansion proofs with cuts, giving translations that are fully canonical in one direction, and partially in the other. Although we do not provide a new method of cut elimination for expansion proofs, we compare the different cut elimination methods given in the above papers, and investigate how they might be implemented in a deep inference proof system. Thus, expansion proofs, in a way, provide an 'off-the-shelf' first-order analogy to atomic flows: they represent certain key information contained in a proof, and can be used to guide normalisation. The theory is not yet fully complete, but the use of expansion proofs provides a more geometric approach to first-order cut elimination for deep inference than the approach described in [Brü06a], and is less particular to one proof system.

Thus, we have two different approaches to Herbrand's theorem: Herbrand proofs and expansion proofs. The first approach is from a more Hilbertian line of proof theory, with links often made to model theory and Gentzen's cut elimination results; the other integrated with newer traditions, such as game semantics and proof nets. However, apart from the definition of new inference rules, no real syntactic innovations are needed to situate both these approaches within deep-inference proof systems, showcasing their capacity to internally describe a wide range of proof theoretic approaches.

The Horizon

While being fairly self-contained, this thesis also represents the first step in an ambitious, somewhat inchoate project. The atomic flow and subatomic logic have shown that the move to deep inference allows us fundamentally new ways to think about the proof theory of a wide array of propositional logics, both individually and collectively. Thus far, the same cannot be said for predicate logics. While a significant portion of deep-inference proof theory has been extended from propositional to first-order logics, including the work presented in this thesis, much of this first-order work has been incremental, rather than revolutionary. In particular, the material connecting deep inference proofs to Herbrand proofs and expansion proofs suggests deep inference as a natural setting for certain branches of first-order proof theory, hopefully widening the appeal of deep-inference proof theory to new proof-theoretic communities. However, it is not claimed that these results represent any sort of Copernican shift, no radically new approach to the subject. Nevertheless, it is hoped that this thesis can act as a launching pad for more provocative and unexpected work.

In collaboration with Guglielmi and Michel Parigot, investigations have been conducted into a first-order proof system with explicit/indicated substitutions over the course of the last few years, with some preliminary and provisional definitions and theorems sketched out. Since, however, we have not yet been able to coalesce some of the more ambitious topics of these discussions into concrete definitions, theorems and proofs, it was decided to leave them out from this thesis. The central idea is that it might be possible and desirable to distinguish and detach two functions of quantifiers in derivations: their role in closing free variables and their role in assigning terms to variables. In particular, it is suggested that the instantiation rule might be decomposed into two types of rule: one choosing a term as a witness for the quantified variable; one removing the quantifier and freeing the variable. These two roles are already distinguished to a certain extent by Herbrand proofs: the prenexification phase removes any locality from the binding of variables, whereas the instantiation phase assigns witnesses.

Two analogies, from very different proof theoretic traditions, might be useful

$$\mathsf{n}{\downarrow}\frac{[\tau \Rightarrow x]A}{\exists x A} \quad \longrightarrow \quad \mathsf{ex}{\downarrow}\frac{\mathsf{ex}{\uparrow}\dfrac{[\tau \Rightarrow x]A}{[\tau \to x]\left(\mathsf{rn}{\downarrow}\dfrac{A}{\exists x A}\right)}}{\exists x A}$$

Figure 6: The instantiation rule decomposed

here. The first is Girard's analysis of intuitionistic implication $A \to B$ as a combination of linear implication and the bang exponential, $!A \multimap B$ [Gir87]. Here, semantic concerns motivated a syntactic innovation that led the discovery to a whole new class of logics. Similarly, it is possible that the addition of explicit/indicated substitutions to classical first-order logic may suggest the definition of new logics, or new proof systems for old logics, perhaps intermediate logics, for which hypersequent calculi have been defined [Avr91; CGT08].

The second is not as obvious, but an example from the proof theory of classical first-order logic: Hilbert's ϵ-calculus [HB74], which has seen a renewal of interest in recent years [AB16; AHW18; BLL18; Min08; MZ06]. The ϵ-calculus was developed as a tool for proving the consistency of arithmetic, as part of Hilbert's program, discussed above. The two ϵ-theorems show that the usual quantifiers can be jettisoned in favour of enriching the term structure, with this 'ϵ-substitution' method allowing for a new consistency proof of first-order arithmetic [Ack40]. This rearrangement of first-order information contained in a proof has significant proof theoretic consequences, and it might be expected that indicated/explicit substitutions might too. Of course, it would likely be hubristic to claim that this project will have a comparable legacy to either of these two innovations, but hopefully the analogies are useful in describing the scope of this more ambitious project.

For now, investigations are proceeding into a new definition of analyticity using indicated substitutions, leading to stronger decomposition theorems. Looking further ahead, it is hoped that the use of indicated substitutions will enable us to fully distinguish between the first-order and propositional aspects of proofs, in a more thorough way than is achieved by the use of expansion proofs. It is expected that extending the rules of passage to substitutions will be key to the endeavour.

The results in this thesis may not seem particularly pertinent to this larger project, but, since the means of normalisation and normal forms are likely to be both extensions of those for classical propositional logic, and comparable to those for classical first-order logic, understanding the theory of normalisation and, in particular, decomposition for these logics will provide a stable platform on which to proceed into more speculative research. Any paradigm shift is preceded by the exhausting of the previous paradigm, finding out where the old approach reaches its explicatory limits, discovering where precisely the patient exhibits morbid symptoms.

Summary

Part I

In Chapter 1, the deep inference formalism of open deduction is presented, and the standard proof systems for classical propositional logic are set out. Some standard theory for propositional logic is then presented—the reduction of structural rules to atomic form, and the isolation of identies and cuts—culminating in a simple and interesting method for cut elimination: the experiments method.

In Chapter 2, the proof theory of propositional logic is developed further: we describe reduction systems for atomic weakening and contraction and show that the barrier to strong normalisation for the contraction system is the existence of cycles in proofs. We then introduce contractive derivations and merge contractions, developing some theory so that we are in a position to present a method for eliminating cycles in proofs by permuting critical merge contractions down the proof.

Part II

In Chapter 3, we introduce first-order classical logic, and give the standard proof systems. We then show that the basic theory of contractive derivations and merge contractions extend to first-order logic in a straightforward way.

In Chapter 4, we present a study of Herbrand's Theorem from the point of view of deep inference. We describe a class of first-order proofs, Herbrand Proofs, and state and prove Herbrand's Theorem for cut-free first-order proofs in open deduction as a transformation into a Herbrand proof. We introduce expansion proofs, and a class of first-order proof that corresponds closely to them: proofs in Herbrand Normal Form. We then show the translations between Expansion Proofs and proofs in Herbrand Normal Form. Finally, we show cut elimination for proofs in Herbrand Normal Form, and how it closely corresponds to cut elimination for expansion proofs with cuts, also known as proof forests. This culminates in a cut elimination procedure for first-order logic via Herbrand's Theorem.

A note on originality and collaboration

Only parts of this thesis are truly original and even less is solely my own work, for the usual expository reasons and so that the fruits of productive collaboration can be shared. The early material on open deduction for propositional logic is a new treatment of standard material, and the idea behind the experiments method was shown to me by Guglielmi, while the formal presentation is essentially my own. Chapter 2 is a result of a long and fruitful collaboration between Aler, Guglielmi and myself, and the cycle removal procedure has taken many forms over the last few years. An early, simpler version can be found in

[ATG18], and contractive derivations can be found in the original version of Aler's PhD thesis under the guise of *generic contractions* [AT16]. I do, however, take credit for reworking these generic contractions as standalone inference rules, *merge contractions*, even if the idea is borrowed from a rather different setting, [Gug07]. Hence the focus on merge contractions in the cycle removal procedure, with extensive theoretical grounding, and a further development of their theory for first-order logic. However the basic idea for cycle removal is not so different as in Aler's thesis, and the presentation in the revised version is more similar to that presented here, but with important differences [AT17].

First-order logic, and Herbrand's theorem in particular was the original focus of my PhD, and thus Part II contains the bulk of my original contributions to deep inference proof theory. Although Brünnler laid exceptional foundations for this work, I believe that the organisation of the first-order material, in particular the centrality of Herbrand's Theorem and its use as a decomposition theorem as a means for cut elimination, is essentially novel, as are the translations between proofs in Herbrand Normal Form and expansion proofs, which have already been published separately [Ral18]. Straßburger has also shown translations between expansion proofs and deep inference proofs for second-order MLL, but I was not aware of this work until we discussed the work that led to the paper in which my own ideas were first published [Str09; Str17b]. The cut elimination section is more of a summary and comparison of the work in [Hei10], [McK13] and [AHW18], as is made clear in the text.

As a totality, it is hoped that this thesis provides a useful overview of the state-of-the-art in the deep inference proof theory of classical logic, and will be both a high-level and technical resource for those new to this field. It is clearly not an exhaustive work, and readers are encouraged to attempt to fill in any lacunae that they encounter and to upgrade conjectures to theorems.

Finally, it should be made unequivocally clear that any errors or misjudgements are, of course, my sole and complete responsibility.

Part I

Propositional Logic

Chapter 1

Open Deduction for Classical Propositional Logic

1.1 Open Deduction

We present the *open deduction* proof formalism [BM08; GGP10], which has sometimes been named the functorial calculus [Gun09]. Inevitably, this presentation of the fundamentals will borrow and differ from previous iterations, especially [GGP10] and [Gun09], but not importantly so.

1.1.1 Prederivations and formulae

In the formalism of open deduction, it is helpful to think of derivations as ontologically prior to formulae, and so we define derivations—or prederivations, to be precise—before we define formulae.

Definition 1.1. Let $\Sigma = (\mathcal{A}, \mathcal{U}, \mathcal{C}_1, \mathcal{C}_2)$ be a signature, with $\mathcal{A}$ a set of atoms, $\mathcal{U}$ a set of units, and $\mathcal{C}_i$ a set of logical operators of arity i, each equipped with an involutive bijection, *negation* commonly denoted by $\bar{\cdot}$. Then a *prederivation* over Σ is inductively defined in the following way:

- All atoms $a \in \mathcal{A}$ and all units $u \in \mathcal{U}$ are prederivations.

If ϕ and ψ are prederivations, then:

- If $\bullet \in \mathcal{C}_1$ then $\bullet\phi$ is a prederivation.
- If $\star \in \mathcal{C}_2$ then $\phi \star \psi$ is a prederivation.
- $\dfrac{\phi}{\psi}$ is a prederivation.

We denote the set of prederivations over Σ by $\mathcal{P}_\Sigma$, omitting the subscript when we are talking about prederivations of an unspecified signature and when otherwise appropriate. We use $\equiv$ for syntactic equality between two prederivations, since we will go on to define an equality relation $=$ on formulae.

Negation is extended to $\mathcal{P}_\Sigma$ in the following way:

- $\overline{\bullet\phi} \equiv \bar{\bullet}\overline{\phi}$
- $\overline{\phi \star \psi} \equiv \overline{\phi}\bar{\star}\overline{\psi}$
- $\overline{\left(\dfrac{\phi}{\psi}\right)} \equiv \dfrac{\overline{\psi}}{\overline{\phi}}$

Note that the extension of negation to prederivations is still an involutive bijection.

Remark 1.2. Since we will only encounter unary and binary operators, we restrict our attention to these, although there is no principled reason to rule out operators of higher arity.

Convention 1.3. Vertical composition of prederivations is associative.

$$\begin{array}{c}\phi\\\hline\psi\\\hline\xi\end{array} :\equiv \frac{\phi}{\left(\dfrac{\psi}{\xi}\right)} \equiv \frac{\left(\dfrac{\phi}{\psi}\right)}{\xi}$$

Definition 1.4. A *formula* is a prederivation with no instances of vertical composition, and we denote the set of formulae over Σ by $\mathcal{F}_\Sigma$, again omitting the subscript when appropriate. We define two functions from $\mathcal{P}$ to $\mathcal{F}$, pr (*premise*) and cn (*conclusion*):

- $\mathsf{pr}(a) \equiv \mathsf{cn}(a) \equiv a$, for $a \in \mathcal{A}$;
- $\mathsf{pr}(u) \equiv \mathsf{cn}(u) \equiv u$, for $u \in \mathcal{U}$;
- $\mathsf{pr}(\bullet\phi) \equiv \bullet(\mathsf{pr}(\phi))$, $\mathsf{cn}(\bullet\phi) \equiv \bullet(\mathsf{cn}(\phi))$, for $\bullet \in \mathcal{C}_1$;
- $\mathsf{pr}(\phi \star \psi) \equiv \mathsf{pr}(\phi) \star \mathsf{pr}(\psi)$, $\mathsf{cn}(\phi \star \psi) \equiv \mathsf{cn}(\phi) \star \mathsf{cn}(\psi)$, for $\star \in \mathcal{C}_2$;
- $\mathsf{pr}\left(\dfrac{\phi}{\psi}\right) \equiv \mathsf{pr}(\phi)$, $\mathsf{cn}\left(\dfrac{\phi}{\psi}\right) \equiv \mathsf{cn}(\psi)$

We can now write a prederivation as $\begin{array}{c}A\\ \phi\,\|\\ B\end{array}$, where $\mathsf{pr}(\phi) \equiv A$ and $\mathsf{cn}(\phi) \equiv B$.

Definition 1.5. The *sections* of a prederivation, $\mathsf{sec}(\phi)$ are defined in the following way. If $A \in \mathsf{sec}(\phi)$, we say A is a *section of* ϕ.

- $\mathsf{sec}(a) = \{a\}$, for $a \in \mathcal{A}$;
- $\mathsf{sec}(u) = \{u\}$, for $u \in \mathcal{U}$;
- $\mathsf{sec}(\bullet\phi) = \{\bullet A \mid A \in \mathsf{sec}(\phi)\}$, for $\bullet \in \mathcal{C}_1$;
- $\mathsf{sec}(\phi \star \psi) = \{A \star B \mid A \in \mathsf{sec}(\phi), B \in \mathsf{sec}(\psi)\}$, for $\star \in \mathcal{C}_2$;
- $\mathsf{sec}\left(\dfrac{\phi}{\psi}\right) = \mathsf{sec}(\phi) \cup \mathsf{sec}(\psi)$

Definition 1.6. Given a signature $\Sigma = (\mathcal{A}, \mathcal{U}, \mathcal{C}_1, \mathcal{C}_2)$ and functions $|\cdot| : \mathcal{A} \to \mathbb{N}, |\cdot| : \mathcal{U} \to \mathbb{N}$, we can define the *size* of a prederivation is a function $|\cdot| : \mathcal{P}_\Sigma \to \mathbb{N}$ defined in the following way:

- The size of atoms and units are already defined by the given functions.
- $\left|\bullet(\phi)\right| = \left|\phi\right| + 1$
- $\left|\phi \star \psi\right| \left|\phi\right| + \left|\psi\right| + 1$
- $\left|\rho \frac{\phi}{\psi}\right| = \left|\phi\right| + \left|\psi\right|$

Definition 1.7. *(Prederivation) contexts with n holes* are certain inductively defined functions $\kappa\overbrace{\{\ \}\ldots\{\ \}}^{n} : \mathcal{P}^n \to \mathcal{P}$, where prederivation contexts with 0 holes are identified with prederivations, and the identity function, denoted $\{\ \}$, is a context with 1 hole. More complex contexts are built up in the same way as prederivations, with application performed by replacing gaps with prederivations. We say that κ is a *formula context* if it contains no instances of vertical composition. We say κ is a *formula-like context* if no hole is in the scope of vertical composition.

The size of a context is calculated in the same fashion as prederivations, with $|\{\ \}| = 0$.

Example 1.8. Let Σ be the signature $(\{a, \bar{a}\}, \{\mathsf{t}, \mathsf{f}\}, \{\Box, \Diamond\}, \{\wedge, \vee\})$. Then

$$\phi \equiv \frac{\bar{a}}{\Box\left(\frac{\mathsf{f}}{\mathsf{t}} \wedge \Diamond\left(\frac{\mathsf{t}}{\mathsf{f}}\right)\right)}$$

is a prederivation over Σ, with $\mathsf{pr}(\phi) \equiv \bar{a}$, $\mathsf{cn}(\phi) \equiv \Box(\mathsf{t} \wedge \Diamond(\mathsf{f}))$ and $\mathsf{sec}(\phi) = \{\bar{a}, \Box(\mathsf{f} \wedge \Diamond\mathsf{t}), \Box(\mathsf{f} \wedge \Diamond\mathsf{f}), \Box(\mathsf{t} \wedge \Diamond\mathsf{t}), \Box(\mathsf{t} \wedge \Diamond\mathsf{f})\}$.

We can also define the prederivation context with 2 holes $\kappa_1\{\ \}\{\ \}$ by:

$$\kappa_1\{\ \}\{\ \} = \frac{\bar{a}}{\Box(\{\ \} \wedge \{\ \})},$$

with $\phi \equiv \kappa_1 \left\{\frac{\mathsf{f}}{\mathsf{t}}\right\}\left\{\Diamond\left(\frac{\mathsf{t}}{\mathsf{f}}\right)\right\}$.

$\kappa_1\{\ \}\{\ \}$ is neither a formula context nor formula-like, $\kappa_2\{\ \} = \Box\left(\frac{\mathsf{f}}{\mathsf{t}} \wedge \{\ \}\right)$ is a formula-like context and $\kappa_3\{\ \}\{\ \} = \Box(\{\ \} \wedge \{\ \})$ is a formula context.

We can also compute the negation of ϕ:

$$\bar{\phi} \equiv \frac{\Diamond\left(\frac{\mathsf{f}}{\mathsf{t}} \vee \Box\left(\frac{\mathsf{t}}{\mathsf{f}}\right)\right)}{a}$$

1.1.2 Categorical Composition of Prederivations

We will now show another way to vertically compose two prederivations, if the conclusion of the first is equal to the premise is the second. We call this categorical composition of prederivations, since this is its denotation in categorical semantics of deep inference proof systems [Hug04; McK06].

Definition 1.9. Let $\phi, \psi \in \mathcal{P}$, with $\mathsf{cn}(\phi) \equiv \mathsf{pr}(\psi)$. Then the *categorical composition* of ϕ and ψ, denoted $\begin{array}{c}\phi\\ \cdots\\ \psi\end{array}$ is defined inductively as follows:

- If $\phi \in \mathcal{A}$, then we define $\begin{array}{c}\phi\\ \cdots\\ \psi\end{array} \equiv \psi$
- If $\phi \equiv \bullet\phi'$, with $\bullet \in \mathcal{C}_1$, then we must have that $\mathsf{cn}(\phi) \equiv \bullet A \equiv \mathsf{pr}(\psi)$, for $A \in \mathcal{F}$. We either have $\psi \equiv \bullet(\psi')$ or $\psi \equiv \dfrac{\psi_1}{\psi_2}$. We deal with each case as below:

$$\begin{array}{c}\phi\\ \cdots\\ \psi\end{array} \equiv \bullet\left(\begin{array}{c}\phi'\\ \cdots\\ \psi'\end{array}\right) \quad \text{or} \quad \begin{array}{c}\phi\\ \cdots\\ \psi\end{array} \equiv \dfrac{\left(\begin{array}{c}\phi\\ \cdots\\ \psi_1\end{array}\right)}{\psi_2}$$

- If $\phi \equiv \phi_1 \star \phi_2$, with $\star \in \mathcal{C}_2$ then we must have that $\mathsf{cn}(\phi) \equiv (A_1 \star A_2) \equiv \mathsf{pr}(\psi)$, for $A_1, A_2 \in \mathcal{F}$. We either have $\psi \equiv \psi_1 \star \psi_2$ or $\psi \equiv \dfrac{\psi_1}{\psi_2}$. We deal with each case as below:

$$\begin{array}{c}\phi\\ \cdots\\ \psi\end{array} \equiv \left(\begin{array}{c}\phi_1\\ \cdots\\ \psi_1\end{array} \star \begin{array}{c}\phi_2\\ \cdots\\ \psi_2\end{array}\right) \quad \text{or} \quad \begin{array}{c}\phi\\ \cdots\\ \psi\end{array} \equiv \dfrac{\left(\begin{array}{c}\phi\\ \cdots\\ \psi_1\end{array}\right)}{\psi_2}$$

- If $\phi \equiv \dfrac{\phi_1}{\phi_2}$, we define

$$\begin{array}{c}\phi\\ \cdots\\ \psi\end{array} \equiv \dfrac{\phi_1}{\left(\begin{array}{c}\phi_2\\ \cdots\\ \psi\end{array}\right)}$$

If ξ is the categorical composition of $\begin{array}{c}A\\ \phi\,\|\\ B\end{array}$ and $\begin{array}{c}B\\ \psi\,\|\\ C\end{array}$, we write

$$\begin{array}{c}A\\ \xi\,\|\\ C\end{array} \equiv \begin{array}{c}A\\ \phi\,\|\\ B\\ \psi\,\|\\ C\end{array}$$

Lemma 1.10. *Categorical composition of prederivations is associative.*

Proof. Let ϕ, ψ and ξ be prederivations. We will prove by induction on $|\phi| + |\xi|$ that:

$$\begin{array}{c}\phi\\ \hdashline \left(\begin{array}{c}\psi\\ \hdashline \xi\end{array}\right)\end{array} \equiv \begin{array}{c}\left(\begin{array}{c}\phi\\ \hdashline \psi\end{array}\right)\\ \hdashline \xi\end{array}$$

- If $\phi \equiv a \in \mathcal{A}$, then clearly

$$\begin{array}{c}a\\ \hdashline \left(\begin{array}{c}\psi\\ \hdashline \xi\end{array}\right)\end{array} \equiv \begin{array}{c}\psi\\ \hdashline \xi\end{array} \equiv \begin{array}{c}\left(\begin{array}{c}a\\ \hdashline \psi\end{array}\right)\\ \hdashline \xi\end{array}$$

  The cases where $\psi \equiv a$ or $\xi \equiv a$ are similar.

- If we have that $\xi \equiv \dfrac{\xi_1}{\xi_2}$, then we have, by the IH

$$\begin{array}{c}\phi\\ \hdashline \left(\begin{array}{c}\psi\\ \hdashline \xi\end{array}\right)\end{array} \equiv \begin{array}{c}\phi\\ \hdashline \left(\begin{array}{c}\psi\\ \hdashline \left(\dfrac{\xi_1}{\xi_2}\right)\end{array}\right)\end{array} \equiv \begin{array}{c}\phi\\ \hdashline \left(\dfrac{\left(\begin{array}{c}\psi\\ \hdashline \xi_1\end{array}\right)}{\xi_2}\right)\end{array} \equiv \dfrac{\left(\begin{array}{c}\phi\\ \hdashline \left(\begin{array}{c}\psi\\ \hdashline \xi_1\end{array}\right)\end{array}\right)}{\xi_2} \equiv \dfrac{\left(\begin{array}{c}\left(\begin{array}{c}\phi\\ \hdashline \psi\end{array}\right)\\ \hdashline \xi_1\end{array}\right)}{\xi_2} \equiv \begin{array}{c}\left(\begin{array}{c}\phi\\ \hdashline \psi\end{array}\right)\\ \hdashline \left(\dfrac{\xi_1}{\xi_2}\right)\end{array}$$

  The case where $\phi \equiv \dfrac{\phi_1}{\phi_2}$ is symmetrical.

- If we have that $\psi \equiv \dfrac{\psi_1}{\psi_2}$, then we have

$$\begin{array}{c}\phi\\ \hdashline \left(\begin{array}{c}\left(\dfrac{\psi_1}{\psi_2}\right)\\ \hdashline \xi\end{array}\right)\end{array} \equiv \begin{array}{c}\phi\\ \hdashline \left(\dfrac{\psi_1}{\left(\begin{array}{c}\psi_2\\ \hdashline \xi\end{array}\right)}\right)\end{array} \equiv \dfrac{\left(\begin{array}{c}\phi\\ \hdashline \psi_1\end{array}\right)}{\left(\begin{array}{c}\psi_2\\ \hdashline \xi\end{array}\right)} \equiv \begin{array}{c}\left(\dfrac{\left(\begin{array}{c}\phi\\ \hdashline \psi_1\end{array}\right)}{\psi_2}\right)\\ \hdashline \xi\end{array} \equiv \begin{array}{c}\left(\begin{array}{c}\phi\\ \hdashline \left(\dfrac{\psi_1}{\psi_2}\right)\end{array}\right)\\ \hdashline \xi\end{array}$$

- $\phi \equiv \bullet\phi'$, $\psi \equiv \bullet\psi'$ and $\xi \equiv \bullet\xi'$ then we have, by the IH, that:

$$\begin{array}{c}\phi\\ \hdashline \left(\begin{array}{c}\psi\\ \hdashline \xi\end{array}\right)\end{array} \equiv \bullet\left(\begin{array}{c}\phi'\\ \hdashline \left(\begin{array}{c}\psi'\\ \hdashline \xi'\end{array}\right)\end{array}\right) \equiv \bullet\left(\begin{array}{c}\left(\begin{array}{c}\phi'\\ \hdashline \psi'\end{array}\right)\\ \hdashline \xi'\end{array}\right) \equiv \begin{array}{c}\left(\begin{array}{c}\phi\\ \hdashline \psi\end{array}\right)\\ \hdashline \xi\end{array}$$

- If we have $\phi \equiv \phi_1 \star \phi_2$, $\psi \equiv \psi_1 \star \psi_2$ and $\xi \equiv \xi_1 \star \xi_2$, then we have, by the IH, that:

$$\begin{array}{c}\phi\\ \hdashline \left(\begin{array}{c}\psi\\ \hdashline \xi\end{array}\right)\end{array} \equiv \left(\begin{array}{c}\phi_1\\ \hdashline \left(\begin{array}{c}\psi_1\\ \hdashline \xi_1\end{array}\right)\end{array} \star \begin{array}{c}\phi_2\\ \hdashline \left(\begin{array}{c}\psi_2\\ \hdashline \xi_2\end{array}\right)\end{array}\right) \equiv \left(\begin{array}{c}\left(\begin{array}{c}\phi_1\\ \hdashline \psi_1\end{array}\right)\\ \hdashline \xi_1\end{array} \star \begin{array}{c}\left(\begin{array}{c}\phi_2\\ \hdashline \psi_2\end{array}\right)\\ \hdashline \xi_2\end{array}\right) \equiv \begin{array}{c}\left(\begin{array}{c}\phi\\ \hdashline \psi\end{array}\right)\\ \hdashline \xi\end{array}$$

Since $\mathsf{cn}(\phi) \equiv \mathsf{pr}(\psi)$ is a precondition for $\begin{array}{c}\phi\\ \hdashline \psi\end{array}$ being well-defined, we have exhausted all the cases. □

Remark 1.11. We can now define a category $\mathcal{C}_\Sigma$, with formula as objects and prederivations from A to B as morphisms. The identity morphism on a formula is the formula itself, treated as a prederivation.

Lemma 1.12. *If* $\phi \equiv \kappa\left\{\frac{\psi}{\xi}\right\}$ *then we can find a unique formula-like context* $\theta\{\ \}$ *and prederivations* ϕ_1, ϕ_2 *s.t.*

$$\phi \equiv \begin{array}{c}\phi_1\\ \hdashline \theta\left\{\dfrac{\mathsf{cn}\,\psi}{\mathsf{pr}\,\xi}\right\}\\ \hdashline \phi_2\end{array}$$

Proof. We proceed by structural induction on κ. If $\kappa = \{\ \}$, we take $\phi_1 \equiv \psi$ and $\phi_2 \equiv \xi$.

If $\kappa\{\ \} = \bullet\kappa'\{\ \}$, then by the inductive hypothesis we can find:

$$\begin{array}{c}\phi_1'\\ \hdashline \theta\left\{\dfrac{\mathsf{cn}\,\psi}{\mathsf{pr}\,\xi}\right\}\\ \hdashline \phi_2'\end{array} \equiv \kappa'\left\{\frac{\psi}{\xi}\right\}$$

We then take $\phi_1 \equiv \bullet\phi_1'$ and $\phi_2 \equiv \bullet\phi_2'$.

If $\kappa\{\ \} = \chi \star \kappa'\{\ \}$, then, similarly to above, we can find appropriate ϕ_1' and ϕ_2' such that we can take $\phi_1 \equiv \chi \star \phi_1'$ and $\phi_2 \equiv \chi \star \phi_2'$.

If $\kappa\{\ \} = \dfrac{\chi}{\kappa'\{\ \}}$, then by the IH we can find ϕ_1' and ϕ_2' as above. We take $\phi_1 \equiv \begin{array}{c}\chi\\ \hdashline \phi_1'\end{array}$ and $\phi_2 \equiv \phi_2'$. If $\kappa\{\ \} = \dfrac{\kappa'\{\ \}}{\chi}$ we take $\phi_1 \equiv \phi_1'$ and $\phi_2 \equiv \begin{array}{c}\phi_2'\\ \hdashline \chi\end{array}$. □

1.1.3 Sequential and Synchronal Form

We now introduce two canonical forms for derivations: sequential and synchronal form.

Sequential form is essentially the *calculus of structures* [Brü06a; BT01] represented in open deduction. One way to think about a prederivation in sequential form is that it requires the vertical order on inference rules to be total. An important use for sequential form is to provide an induction measure for open deduction proofs, which are not always straightforward for two-dimensional proofs.

Synchronal form, originally called *Formalism A* [Gug04], can be thought of as a form where any non-essential vertical ordering between inference rules is eliminated. Since the synchronisation reduction is confluent, a prederivation in synchronal form is a canonical form for a class of prederivations [GGP10].

Definition 1.13. We write $\equiv\dfrac{\phi}{\psi}$ if $\mathsf{cn}(\phi) \equiv \mathsf{pr}(\psi)$.

Definition 1.14. A prederivation ϕ is in *sequential form* if, for every possible $\kappa\{\ \}, \psi, \xi$ s.t $\phi \equiv \kappa\left\{\dfrac{\psi}{\xi}\right\}$, $\kappa'\{\ \}$ as in Lemma 1.12 is a formula context. Equivalently, there are κ_i, A_i, B_i s.t. :

$$\phi \equiv \begin{array}{c} A \\ \hline \kappa_1\left\{\dfrac{A_1}{B_1}\right\} \\ \hline \vdots \\ \hline \kappa_n\left\{\dfrac{A_n}{B_n}\right\} \\ \hline B \end{array}$$

A prederivation is in *synchronal form* if, for every possible $\kappa\{\ \}, \psi, \xi$ s.t $\phi \equiv \kappa\left\{\dfrac{\psi}{\xi}\right\}$, $\mathsf{cn}(\psi) \not\equiv \mathsf{pr}(\xi)$.

Definition 1.15. A *reduction rule* r is a partial function $r : \mathcal{P}_\Sigma \to \mathcal{P}$ s.t. if $r(\phi') \equiv \psi'$, then ϕ' and ψ' have the same premise and conclusion. We write $r : \phi' \to \psi'$.

For every reduction rule $r : \phi' \to \psi'$ we define the reduction $\to_r$ such that $\phi \to_r \psi$ iff ϕ' is a subprederivation of ϕ and ψ is obtained from ϕ by replacing ϕ' by ψ'.

We call a finite set R of reduction rules a *rewriting system*. We write $\phi \to_R \psi$ if there is $r \in R$ s.t. $\phi \to_r \psi$. The reflexive transitive closure of $\to_r$ is written $\to_r^*$. Given a set S of prederivations, we say that rewriting system R is *terminating on S* if there is no infinite chain $\phi \to_{r_1} \phi_1 \to_{r_2} \dots$ with $r_i \in R$ for any $\phi \in S$. If ϕ is s.t. there is no ψ with $\phi \to_r \psi$ we say ϕ is *normal* for R, or that ϕ is in *normal form* w.r.t. R.

Definition 1.16. We define the rewriting system Seq as containing the following two rewrites S_1 and S_2, where κ is a formula-like context:

$$\equiv\dfrac{\kappa\left\{\dfrac{A_1}{B_1}\right\}\{A_2\}}{\kappa\{B_1\}\left\{\dfrac{A_2}{B_2}\right\}} \quad \xleftarrow{S_1} \quad \kappa\left\{\dfrac{A_1}{B_1}\right\}\left\{\dfrac{A_2}{B_2}\right\} \quad \xrightarrow{S_2} \quad \equiv\dfrac{\kappa\{A_1\}\left\{\dfrac{A_2}{B_2}\right\}}{\kappa\left\{\dfrac{A_1}{B_1}\right\}\{B_2\}}$$

We define the rewriting system Syn as containing the single rewrite S_3:

$$\equiv\frac{\left(\begin{array}{c}A\\ \phi\,\Vert\\ B\end{array}\right)}{\left(\begin{array}{c}B\\ \psi\,\Vert\\ C\end{array}\right)} \quad\xrightarrow{S_3}\quad \begin{array}{c}A\\ \phi\,\Vert\\ B\\ \psi\,\Vert\\ C\end{array}$$

If $\phi \longrightarrow^*_{\mathsf{Seq}} \psi$ with ψ in sequential form, we say that ψ is a *sequentialisation* of ϕ. If $\phi \longrightarrow^*_{\mathsf{Syn}} \psi$ with ψ in synchronal form, we say that ψ is a *synchronisation* of ϕ.

Proposition 1.17. Seq *and* Syn *are terminating on* $\mathcal{P}_\Sigma$ *for any signature* Σ.

Remark 1.18. S_3 is the left inverse of both S_1 and S_2.

Proposition 1.19. ϕ *is in normal form w.r.t.* Seq *iff* ϕ *is in sequential form.* ϕ *is in normal form w.r.t.* Syn *iff* ϕ *is in synchronal form.*

Proposition 1.20 (Confluence of Syn)**.** *Assume* $\phi \longrightarrow^*_{Syn} \psi_1$ *and* $\phi \longrightarrow^*_{Syn} \psi_2$*, with* ψ_1 *and* ψ_2 *in normal form w.r.t.* Syn*. Then* $\psi_1 \equiv \psi_2$*.*

Proof. Since Syn is clearly terminating, we need only prove local confluence. Let ϕ be a derivation with two Syn redexes

$$\phi \equiv \kappa_1\left\{\equiv\frac{\begin{array}{c}A_1\\ \psi_1\,\Vert\\ B_1\end{array}}{\begin{array}{c}B_1\\ \xi_1\,\Vert\\ C_1\end{array}}\right\} \quad\text{and}\quad \phi \equiv \kappa_2\left\{\equiv\frac{\begin{array}{c}A_2\\ \psi_2\,\Vert\\ B_2\end{array}}{\begin{array}{c}B_2\\ \xi_2\,\Vert\\ C_2\end{array}}\right\}$$

Either the redexes overlap or they don't:

$$\phi \equiv \kappa_3\left\{\equiv\frac{\begin{array}{c}A_1\\ \psi_1\,\Vert\\ B_1\end{array}}{\begin{array}{c}B_1\\ \xi_1\,\Vert\\ C_1\end{array}}\right\}\left\{\equiv\frac{\begin{array}{c}A_2\\ \psi_2\,\Vert\\ B_2\end{array}}{\begin{array}{c}B_2\\ \xi_2\,\Vert\\ C_2\end{array}}\right\} \quad\text{or}\quad \phi \equiv \kappa_1\left\{\equiv\frac{\kappa_2'\left\{\equiv\frac{\begin{array}{c}A_2\\ \psi_2\,\Vert\\ B_2\end{array}}{\begin{array}{c}B_2\\ \xi_2\,\Vert\\ C_2\end{array}}\right\}}{\begin{array}{c}B_1\\ \xi_1\,\Vert\\ C_1\end{array}}\right\}$$

If they do not, the first case, then clearly we have local confluence. If they do, then by Lemma 1.12, we can rewrite ϕ in the following way, which clearly rewrites to the same prederivation whichever rewrite is performed first:

$$\phi \equiv \kappa_1 \left\{ \begin{array}{c} \kappa_3 \left\{ \begin{array}{c} A_2 \\ \| \\ B_2 \end{array} \right\} \\ \cdots\cdots\cdots \\ \theta \left\{ \equiv \dfrac{B_2}{B_2} \right\} \\ \cdots\cdots\cdots \\ \kappa_4 \left\{ \begin{array}{c} B_2 \\ \| \\ C_2 \end{array} \right\} \\ \hline \equiv \quad B_1 \\ \xi_1 \, \| \\ C_1 \end{array} \right\} \xrightarrow{(S_3)^2} \phi \equiv \kappa_1 \left\{ \begin{array}{c} \kappa_3 \left\{ \begin{array}{c} A_2 \\ \| \\ B_2 \end{array} \right\} \\ \cdots\cdots\cdots \\ \theta \{B_2\} \\ \cdots\cdots\cdots \\ \kappa_4 \left\{ \begin{array}{c} B_2 \\ \| \\ C_2 \end{array} \right\} \\ \cdots\cdots\cdots \\ B_1 \\ \xi_1 \, \| \\ C_1 \end{array} \right\}$$

□

1.1.4 Proof Systems and Derivations

Definition 1.21. An *equality relation*, $=$, for Σ, is an equivalence relation on $\mathcal{F}$; an inference rule ρ, for Σ, is a relation on $\mathcal{F}$. If $(A,B) \in =$, then we write $A = B$ and if $(A,B) \in \rho$, then $\dfrac{A}{B}$ is an *instance* of ρ and we write $\rho\,\dfrac{A}{B}$.

Remark 1.22. Inference rules are often defined in pairs: *up-rules* and *down-rules*. Rules with the same name but opposite arrows, e.g. $\rho\downarrow$ and $\rho\uparrow$, are *dual*: $\rho\!\downarrow\!\dfrac{A}{B}$ is an instance of $\rho\downarrow$ iff $\rho\!\uparrow\!\dfrac{\bar{B}}{\bar{A}}$ is an instance of $\rho\uparrow$.

Definition 1.23. A *proof system* S for Σ is a finite set of inference rules for Σ that necessarily includes an equality relation $=$ for Σ.

Remark 1.24. For proof complexity purposes, it is important to stipulate that inference rules are polynomially verifiable in the size of the premise and conclusion formulae. However since this condition is not used in this thesis, we omit it from the definition.

Definition 1.25. A derivation in S or an S-*derivation*, where S is a proof system for Σ, is a prederivation over Σ where every instance of vertical composition $\dfrac{\psi}{\chi}$ is such that $\dfrac{\mathsf{cn}\,\psi}{\mathsf{pr}\,\chi}$ is an instance of ρ, for $\rho \in \mathsf{S}$. We write $\begin{array}{c} A \\ \phi \| \mathsf{S} \\ B \end{array}$ or $\begin{array}{c} A \\ \phi \| \{\rho_1,\ldots,\rho_n\} \\ B \end{array}$ if ϕ is an S-derivation, where $\mathsf{S} = \{\rho_1, \ldots, \rho_n, =\}$.

Definition 1.26. If S is a proof system with a *true* unit, i.e. $\mathsf{t} \in \mathcal{U}$, and ϕ is an S derivation with $\mathsf{pr}(\phi) \equiv \mathsf{t}$ and $\mathsf{cn}(A)$, then we say ϕ is a *proof* of A, and we write $\begin{array}{c} \phi \rVert \mathsf{S} \\ A \end{array}$.

Remark 1.27. Since we are sticking to classical logic in this thesis, we will not worry too much about what counts as a 'true' unit.

Definition 1.28. An inference rule ρ is *admissible* for a proof system S if for every proof $\begin{array}{c}\phi \,\Vert\, \mathsf{S}\cup\{\rho\} \\ A\end{array}$ there is a proof $\begin{array}{c}\phi' \,\Vert\, \mathsf{S} \\ A\end{array}$.

Definition 1.29. An inference rule ρ is *derivable* for a proof system S if for every instance $\rho\,\frac{A}{B}$ of ρ there is a derivation $\begin{array}{c}A \\ \phi \,\Vert\, \mathsf{S} \\ B\end{array}$.

Observation 1.30. If If ρ is derivable for S, then ρ is admissible for S. If ρ is admissible (derivable) for S and $\mathsf{S} \subseteq \mathsf{S}'$, then ρ is admissible (derivable) for S'.

Definition 1.31. Proof systems S and S' are *equivalent* if, for every formula A, there is a proof $\begin{array}{c}\phi \,\Vert\, \mathsf{S} \\ A\end{array}$ iff there is a proof $\begin{array}{c}\phi \,\Vert\, \mathsf{S}' \\ A\end{array}$.

Definition 1.32. Proof systems S and S' are *strongly equivalent* if, for all formulae A and B, there is a derivation $\begin{array}{c}A \\ \phi \,\Vert\, \mathsf{S} \\ B\end{array}$ iff there is a derivation $\begin{array}{c}A \\ \phi \,\Vert\, \mathsf{S}' \\ B\end{array}$.

Observation 1.33. S and S' are equivalent iff every rule $\rho \in \mathsf{S}$ is admissible for S' and vice versa. S and S' are strongly equivalent iff every rule $\rho \in \mathsf{S}$ is derivable for S' and vice versa.

1.2 Proof Systems for Classical Propositional Logic

We are now in a position to define the standard open deduction proof systems for classical propositional logic (CPL). Since the development of the atomic flow [GGP10; Gun09], SKS and its cut-free subsystem KS have been central to the deep inference proof theory of CPL.

Definition 1.34. We define the signature for classical propositional logic $\Sigma_0 = (\mathcal{A}_0, \{\mathsf{t},\mathsf{f}\}, \emptyset, \{\wedge,\vee\})$, where $\mathcal{A}_0 = \{a, \bar{a}, b, \bar{b}, c, \bar{c} \ldots\}$ is an inexhaustible supply of positive and negative atoms, $\bar{\mathsf{t}} = \mathsf{f}, \bar{\mathsf{t}} = \mathsf{f}$ and $\bar{\wedge} = \vee, \bar{\vee} = \wedge$.

Definition 1.35. The *size* of a CPL prederivation is defined by setting $|a| = |\mathsf{t}| = |\mathsf{f}| = 1$ for every $a \in \mathcal{A}_0$.

1.2.1 SKS

Definition 1.36. The proof system SKS for Σ_0, consists of the following inference rules [Brü03a; GGP10]:

- The *structural* rules:

$$
\textit{atomic}\left\{\begin{array}{ccc}
\mathsf{ai}{\downarrow}\,\dfrac{\mathsf{t}}{a \vee \bar{a}} & \mathsf{ac}{\downarrow}\,\dfrac{a \vee a}{a} & \mathsf{aw}{\downarrow}\,\dfrac{\mathsf{f}}{a} \\
\textit{identity} & \textit{contraction} & \textit{weakening} \\
 & & \\
\mathsf{ai}{\uparrow}\,\dfrac{a \wedge \bar{a}}{\mathsf{f}} & \mathsf{ac}{\uparrow}\,\dfrac{a}{a \wedge a} & \mathsf{aw}{\uparrow}\,\dfrac{a}{\mathsf{t}} \\
\textit{cut} & \textit{cocontraction} & \textit{coweakening}
\end{array}\right\}
$$

- The *logical* rules:

$$
\begin{array}{cc}
\mathsf{s}\,\dfrac{A \wedge (B \vee C)}{(A \wedge B) \vee C} & \mathsf{m}\,\dfrac{(A \wedge B) \vee (C \wedge D)}{(A \vee C) \wedge (B \vee D)} \\
\textit{switch} & \textit{medial}
\end{array}
$$

- The equality relation is the minimal equivalence relation generated by the following equalities:

$$
\begin{array}{cc}
A \wedge \mathsf{t} = A & A \vee \mathsf{f} = A \\
\mathsf{t} \vee \mathsf{t} = \mathsf{t} & \mathsf{f} \wedge \mathsf{f} = \mathsf{f} \\
A \wedge B = B \wedge A & A \vee B = B \vee A \\
A \wedge (B \wedge C) = (A \wedge B) \wedge C & A \vee (B \vee C) = (A \vee B) \vee C
\end{array}
$$

Example 1.37. Below is a derivation in SKS, with premise t and conclusion $\bar{a} \vee a$:

$$
\mathsf{s}\,\dfrac{
= \dfrac{
= \dfrac{\mathsf{t}}{
\boxed{= \dfrac{= \dfrac{\mathsf{t}}{\boxed{\mathsf{ai}{\downarrow}\,\dfrac{\mathsf{t}}{\bar{a} \vee a}} \vee \boxed{\mathsf{aw}{\downarrow}\,\dfrac{\mathsf{f}}{a}}}}{\bar{a} \vee \boxed{\mathsf{ac}{\downarrow}\,\dfrac{a \vee a}{a}}}}
\wedge
\boxed{\mathsf{ai}{\downarrow}\,\dfrac{\mathsf{t}}{\boxed{\mathsf{ac}{\uparrow}\,\dfrac{\bar{a}}{\bar{a} \wedge \boxed{\mathsf{aw}{\uparrow}\,\dfrac{\bar{a}}{\mathsf{t}}}}} \vee a}}
}}{(\bar{a} \vee a) \wedge (\bar{a} \vee a)}
}{
\bar{a} \vee \boxed{\mathsf{s}\,\dfrac{a \wedge (\bar{a} \vee a)}{\boxed{= \dfrac{\boxed{\mathsf{ai}{\uparrow}\,\dfrac{a \wedge \bar{a}}{\mathsf{f}}} \vee a}{a}}}}
}
$$

Often, we employ a few simple abbreviations to minimize size and aid readability. The below proof is identical to the one above except that it employs some of these abbreviations: two switches are compressed into one 2s rule; certain boxes are removed; in the conclusion we leave out parentheses from the disjunction, due to associativity; and equality rules on units are omitted:

$$
= \dfrac{
\boxed{= \dfrac{\mathsf{ai}\!\downarrow \dfrac{\mathsf{t}}{\bar{a} \vee a} \vee \mathsf{aw}\!\downarrow \dfrac{\mathsf{f}}{a}}{\bar{a} \vee \mathsf{ac}\!\downarrow \dfrac{a \vee a}{a}}}
\wedge
\boxed{\mathsf{ai}\!\downarrow \dfrac{\mathsf{t}}{\mathsf{ac}\!\uparrow \dfrac{\bar{a}}{\bar{a} \wedge \mathsf{aw}\!\uparrow \dfrac{\bar{a}}{\mathsf{t}}} \vee a}}
}{
\mathsf{2s} \dfrac{(\bar{a} \vee a) \wedge (\bar{a} \vee a)}{\bar{a} \vee \mathsf{ai}\!\uparrow \dfrac{a \wedge \bar{a}}{\mathsf{f}} \vee a}
}
$$

1.2.2 Other proof systems for CPL

The proof SKSg differs from SKS in two ways: firstly, the structural rules are *generic* not atomic and, secondly, there is no medial, since it is derivable for $\{\mathsf{w}\!\downarrow, \mathsf{c}\!\downarrow\}$ (or $\{\mathsf{w}\!\uparrow, \mathsf{c}\!\uparrow\}$ for that matter).

Definition 1.38. SKSg is defined to be the following inference rules, plus the switch rule.

$$
\begin{array}{ccc}
\mathsf{i}\!\downarrow \dfrac{\mathsf{t}}{A \vee \bar{A}} & \mathsf{c}\!\downarrow \dfrac{A \vee A}{A} & \mathsf{w}\!\downarrow \dfrac{\mathsf{f}}{A} \\
\textit{identity} & \textit{contraction} & \textit{weakening} \\[1em]
\mathsf{i}\!\uparrow \dfrac{A \wedge \bar{A}}{\mathsf{f}} & \mathsf{c}\!\uparrow \dfrac{A}{A \wedge A} & \mathsf{w}\!\uparrow \dfrac{A}{\mathsf{t}} \\
\textit{cut} & \textit{cocontraction} & \textit{coweakening}
\end{array} ,
$$

The equality relation is the same as for SKS.

We now repeat some simple lemmas that establish the strong equivalence of SKS and SKSg [BT01; Brü03a].

Proposition 1.39. *The medial rule* m *is derivable for* $\{\mathsf{w}\!\downarrow, \mathsf{c}\!\downarrow\}$.

Proof. We can replace every instance of medial

$$
\mathsf{m} \dfrac{(A \wedge B) \vee (C \wedge D)}{(A \vee C) \wedge (B \vee D)}
$$

with the following derivation

$$
\mathsf{c}\!\downarrow \dfrac{
\left(\boxed{= \dfrac{A}{A \vee \mathsf{w}\!\downarrow \dfrac{\mathsf{f}}{C}}} \wedge \boxed{= \dfrac{B}{B \vee \mathsf{w}\!\downarrow \dfrac{\mathsf{f}}{D}}} \right)
\vee
\left(\boxed{= \dfrac{C}{\mathsf{w}\!\downarrow \dfrac{\mathsf{f}}{A} \vee C}} \wedge \boxed{= \dfrac{D}{\mathsf{w}\!\downarrow \dfrac{\mathsf{f}}{B} \vee D}} \right)
}{(A \vee C) \wedge (B \vee D)}
$$

. ∎

Lemma 1.40. *The contraction rule* $\mathsf{c}\!\downarrow$ *is derivable for* $\{\mathsf{ac}\!\downarrow, \mathsf{m}\}$, $\mathsf{i}\!\downarrow$ *is derivable for* $\{\mathsf{ac}\!\downarrow, \mathsf{s}\}$ *and* $\mathsf{w}\!\downarrow$ *is derivable for* $\{\mathsf{aw}\!\downarrow\}$.

Proof. We show the proof for contraction, the proofs for the other rules are similar. We show that there is always a derivation $\begin{array}{c} A \vee A \\ \Vert\{\mathsf{ac}\downarrow,\mathsf{m}\} \\ A \end{array}$, by considering four cases:

1. If $A \equiv \mathsf{t}$ or $A = \mathsf{f}$, then we have $= \dfrac{\mathsf{t} \vee \mathsf{t}}{\mathsf{t}}$ or $= \dfrac{\mathsf{f} \vee \mathsf{f}}{\mathsf{f}}$.

2. If $A \equiv a$, then we have $\mathsf{ac}\!\downarrow \dfrac{a \vee a}{a}$.

3. If $A \equiv A_1 \vee A_2$ we have:

$$
= \frac{(A_1 \vee A_2) \vee (A_1 \vee A_2)}{\boxed{\begin{array}{c} A_1 \vee A_1 \\ \Vert\{\mathsf{ac}\downarrow,\mathsf{m}\} \\ A_1 \end{array}} \vee \boxed{\begin{array}{c} A_2 \vee A_2 \\ \Vert\{\mathsf{ac}\downarrow,\mathsf{m}\} \\ A_2 \end{array}}}
$$

4. If $A \equiv A_1 \wedge A_2$ we have:

$$
\mathsf{m} \frac{(A_1 \wedge A_2) \vee (A_1 \wedge A_2)}{\boxed{\begin{array}{c} A_1 \vee A_1 \\ \Vert\{\mathsf{ac}\downarrow,\mathsf{m}\} \\ A_1 \end{array}} \wedge \boxed{\begin{array}{c} A_2 \vee A_2 \\ \Vert\{\mathsf{ac}\downarrow,\mathsf{m}\} \\ A_2 \end{array}}}
$$

Since $|A_1|, |A_2| < |A|$, the size of contracted formulae is suitable as an induction measure. □

Lemma 1.41. *The cocontraction rule* $\mathsf{c}\!\uparrow$ *is derivable for* $\{\mathsf{ac}\!\uparrow, \mathsf{m}\}$, *the cut rule* $\mathsf{i}\!\uparrow$ *is derivable for* $\{\mathsf{ac}\!\uparrow, \mathsf{s}\}$ *and the coweakening rule* $\mathsf{w}\!\uparrow$ *is derivable for* $\{\mathsf{aw}\!\uparrow\}$.

Proof. Dual to the above. □

Proposition 1.42. *The atomic system* SKS *and the generic system* SKSg *are strongly equivalent.*

Proof. As shown above, the medial rule is derivable for SKSg, and clearly the atomic rules are derivable from the generic rules. The lemmas above show that the generic structural rules of SKSg are derivable in SKS. □

Definition 1.43. KS and KSg are, respectively, the systems SKS and SKSg without the three 'up'-rules.

KS and KSg are the standard cut-free/analytic systems for CPL. Thus, cut elimination for SKS is best thought of as showing that the 'up-rules' are admissible for KS.

$$\mathsf{SKS}\quad \mathsf{ai}\!\uparrow\frac{a \wedge \bar{a}}{\mathsf{f}} \qquad \mathsf{ac}\!\uparrow\frac{a}{a \wedge a} \qquad \mathsf{aw}\!\uparrow\frac{a}{\mathsf{t}}$$

$$\mathsf{KS}\quad \mathsf{ai}\!\downarrow\frac{\mathsf{t}}{a \vee \bar{a}} \qquad \mathsf{ac}\!\downarrow\frac{a \vee a}{a} \qquad \mathsf{aw}\!\downarrow\frac{\mathsf{f}}{a} \qquad \mathsf{s}\frac{A \wedge [B \vee C]}{(A \wedge B) \vee C} \qquad \mathsf{m}\frac{(A \wedge B) \vee (C \wedge D)}{[A \vee C] \wedge [B \vee D]}$$

$$\mathsf{SKSg}\quad \mathsf{i}\!\uparrow\frac{A \wedge \bar{A}}{\mathsf{f}} \qquad \mathsf{c}\!\uparrow\frac{A}{A \wedge A} \qquad \mathsf{w}\!\uparrow\frac{A}{\mathsf{t}}$$

$$\mathsf{KSg}\quad \mathsf{i}\!\downarrow\frac{\mathsf{t}}{A \vee \bar{A}} \qquad \mathsf{c}\!\downarrow\frac{A \vee A}{A} \qquad \mathsf{w}\!\downarrow\frac{\mathsf{f}}{A} \qquad \mathsf{s}\frac{A \wedge [B \vee C]}{(A \wedge B) \vee C}$$

Figure 1.1: CPL proof systems SKS, KS, SKSg and KSg

Remark 1.44. Raymond Smullyan defined an analytic proof system to be one where every rule satisfies the subformula property: that the formulae in the premise of an inference rule are subformulae of those in the conclusion [Smu65; Smu68]. This property strictly limits the width of the proof search space, and is closely linked to completeness. In the sequent calculus, the only rule without this property is the cut rule, giving us an identification between cut-freeness and analyticity. However, as discussed in a recent paper by Guglielmi and Bruscoli, this identification breaks down for deep inference: KS does not enjoy anything resembling the subformula property [BG16]. We therefore have two options for the notion in deep inference: directly define analyticity as cut-freeness, or redefine it in the spirit of Smullyan's original definition. In the paper, the authors explore the latter option. Here, we identify analyticity with cut-freeness.

In terms of proof complexity, KS polynomially simulates cut-free Gentzen systems, e.g. LK^- (but not vice versa) while SKS is polynomially equivalent to Gentzen systems with cut LK [BG09; Das15]. It is an open question whether KS polynomially simulates SKS, although it is known that the system $\mathsf{KS} \cup \{\mathsf{ac}\!\uparrow\}$ quasi-polynomially simulates SKS [Bru+16; Das15; Jeř08]. Thus, from the point of view of proof complexity, there is a reasonable but not exact analogy between 'cut-freeness' in Gentzen and deep inference systems.

Theorem 1.45 (Cut Elimination for CPL)**.** KS *and* SKS *are equivalent (but not strongly equivalent).*

In practice, this theorem reduces to the admissibility of just the $\mathsf{ai}\!\uparrow$ rule.

Proposition 1.46. *Each up-rule* $\rho\uparrow \in \mathsf{SKS}$ *or* SKSg *is derivable for* $\{\rho\downarrow, \mathsf{ai}\!\uparrow, \mathsf{ai}\!\downarrow, \mathsf{s}\}$*:*

Proof.

$$\rho\!\uparrow\frac{A}{B} \quad\longrightarrow\quad \cfrac{\mathsf{s}\cfrac{\cfrac{A}{A \wedge \boxed{\mathsf{i}\!\downarrow\cfrac{\mathsf{t}}{\rho\!\downarrow\cfrac{\bar{B}}{\bar{A}} \vee B}}}}{\mathsf{i}\!\uparrow\cfrac{A \wedge \bar{A}}{\mathsf{f}} \vee B}}{B}$$

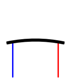
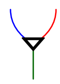
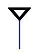
As $\mathsf{i}\!\downarrow$ is derivable for $\{\mathsf{ai}\!\downarrow, \mathsf{s}\}$, and $\mathsf{i}\!\uparrow$ for $\{\mathsf{ai}\!\uparrow, \mathsf{s}\}$ this proves the proposition. □

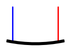
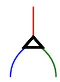
1.3 The Atomic Flow

> "In fact it is not excessive to say that a logic is essentially a set of structural rules!"
>
> — J. Y. Girard [Gir89]

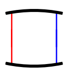
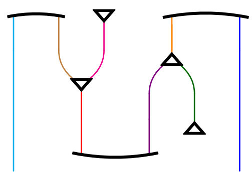
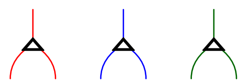
We now introduce the atomic flow, a geometric invariant of proofs in SKS and its subsystems.

We will spare the reader a detailed technical introduction to atomic flows here, but many can be found elsewhere[Das15; GGS10; Gun09]. At a high level, an atomic flow is a directed acyclic graph built up from six types of vertices (corresponding to the six atomic structural rules) which can be equipped with a polarity assignment (but need not be [Das15]). The edges trace the atoms in an SKS derivation, and vertical and horizontal composition of flows correspond to composition of derivations.

$$\mathsf{ai}{\downarrow}\,\frac{\mathsf{t}}{a\vee\bar a}$$  $$\mathsf{ac}{\downarrow}\,\frac{a\vee a}{a}$$  $$\mathsf{aw}{\downarrow}\,\frac{\mathsf{f}}{a}$$ 

$$\mathsf{ai}{\uparrow}\,\frac{a\wedge\bar a}{\mathsf{f}}$$  $$\mathsf{ac}{\uparrow}\,\frac{a}{a\wedge a}$$  $$\mathsf{aw}{\uparrow}\,\frac{a}{\mathsf{t}}$$ 

$$\mathsf{s}\,\cfrac{\mathsf{s}\,\cfrac{\mathsf{m}\,\cfrac{\cfrac{\mathsf{t}}{a\vee\bar a}}{(a\vee\mathsf{t})\wedge(\mathsf{t}\vee\bar a)}}{(a\vee\mathsf{t})\wedge\bar a}}{\cfrac{a\wedge\bar a}{\mathsf{f}}\vee\mathsf{t}}
\qquad
=\cfrac{=\cfrac{\cfrac{\mathsf{t}}{\bar a\vee a}\vee\cfrac{\mathsf{f}}{a}}{\bar a\vee\cfrac{a\vee a}{a}}\wedge\cfrac{\mathsf{t}}{\cfrac{\bar a}{\bar a\wedge\cfrac{\bar a}{\mathsf{t}}}\vee a}}{\mathsf{s}^2\,\cfrac{(\bar a\vee a)\wedge(\bar a\vee a)}{\bar a\vee\cfrac{a\wedge\bar a}{\mathsf{f}}\vee a}}
\qquad
\mathsf{m}\,\cfrac{\cfrac{a}{a\wedge a}\vee\cfrac{b}{b\wedge b}}{(a\vee b)\wedge(a\vee b)}\wedge\cfrac{a}{a\wedge a}$$

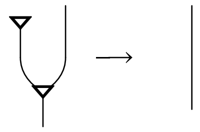
  

A central use of atomic flows is to define rewrite rules for SKS derivations. While one could in theory make do without the atomic flow to define rewriting systems on SKS derivations, in practice they allow for a much simpler and more elegant theory.

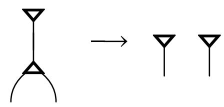
A simple rewriting result for SKS that is made clear by the use of atomic flows is that (co)weakenings can be permuted to the bottom (top) of a derivation using the rules defined below [Das15].

Definition 1.47. We define the following reduction rules for SKS:

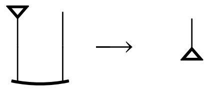
$$\mathsf{aw}{\downarrow}-\mathsf{ac}{\downarrow}:\quad {\scriptstyle \mathsf{ac}\downarrow}\frac{{\scriptstyle \mathsf{aw}\downarrow}\dfrac{\mathsf{f}}{a}\vee a}{a} \longrightarrow {\scriptstyle =}\frac{\mathsf{f}\vee a}{a}$$



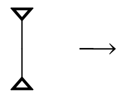
$$\mathsf{aw}{\downarrow}-\mathsf{ac}{\uparrow}:\quad {\scriptstyle \mathsf{ac}\uparrow}\frac{{\scriptstyle \mathsf{aw}\downarrow}\dfrac{\mathsf{f}}{a}}{a\wedge a} \longrightarrow {\scriptstyle \mathsf{aw}\downarrow}\frac{\mathsf{f}}{a}\wedge{\scriptstyle \mathsf{aw}\downarrow}\frac{\mathsf{f}}{a}$$



$$\mathsf{aw}{\downarrow}-\mathsf{ai}{\uparrow}:\quad {\scriptstyle \mathsf{ai}\uparrow}\frac{{\scriptstyle \mathsf{aw}\downarrow}\dfrac{\mathsf{f}}{a}\wedge\bar{a}}{\mathsf{f}} \longrightarrow {\scriptstyle =}\frac{\mathsf{f}\wedge{\scriptstyle \mathsf{aw}\uparrow}\dfrac{\bar{a}}{\mathsf{t}}}{\mathsf{f}}$$

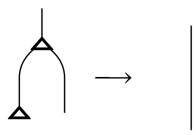


$$\mathsf{aw}{\downarrow}-\mathsf{aw}{\uparrow}:\quad \begin{array}{c}{\scriptstyle \mathsf{aw}\downarrow}\dfrac{\mathsf{f}}{a}\\[-2pt]{\scriptstyle \mathsf{aw}\uparrow}\dfrac{}{\mathsf{t}}\end{array} \longrightarrow \begin{array}{c}{\scriptstyle =}\dfrac{\mathsf{f}}{\mathsf{f}\wedge(\mathsf{f}\vee\mathsf{t})}\\[-2pt]{\scriptstyle \mathsf{s}}\dfrac{}{(\mathsf{f}\wedge\mathsf{f})\vee\mathsf{t}}\\[-2pt]{\scriptstyle =}\dfrac{}{\mathsf{t}}\end{array}$$

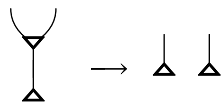


And their duals:

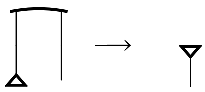
$$\mathsf{ac}{\uparrow}-\mathsf{aw}{\uparrow}:\quad {\scriptstyle \mathsf{ac}\uparrow}\frac{a}{{\scriptstyle \mathsf{aw}\uparrow}\dfrac{a}{\mathsf{t}}\wedge a} \longrightarrow {\scriptstyle =}\frac{a}{\mathsf{t}\wedge a}$$



$$\mathsf{ac}{\downarrow}-\mathsf{aw}{\uparrow}:\quad \begin{array}{c}{\scriptstyle \mathsf{ac}\downarrow}\dfrac{a\vee a}{a}\\[-2pt]{\scriptstyle \mathsf{aw}\uparrow}\dfrac{}{\mathsf{t}}\end{array} \longrightarrow {\scriptstyle \mathsf{aw}\uparrow}\frac{a}{\mathsf{t}}\vee{\scriptstyle \mathsf{aw}\uparrow}\frac{a}{\mathsf{t}}$$



$$\mathsf{ai}{\downarrow}-\mathsf{aw}{\uparrow}:\quad {\scriptstyle \mathsf{ai}\downarrow}\frac{\mathsf{t}}{{\scriptstyle \mathsf{aw}\uparrow}\dfrac{a}{\mathsf{t}}\vee\bar{a}} \longrightarrow {\scriptstyle =}\frac{\mathsf{t}}{\mathsf{t}\vee{\scriptstyle \mathsf{aw}\downarrow}\dfrac{\mathsf{f}}{\bar{a}}}$$



And the trivial reductions:

$$\mathsf{aw}{\downarrow}-\rho_K:\quad \rho\frac{K\left\{{\scriptstyle \mathsf{aw}\downarrow}\dfrac{\mathsf{f}}{a}\right\}}{K'\{a\}} \longrightarrow \rho\frac{K\{\mathsf{f}\}}{K'\left\{{\scriptstyle \mathsf{aw}\downarrow}\dfrac{\mathsf{f}}{a}\right\}}$$

$$\rho_K-\mathsf{aw}{\uparrow}:\quad \rho\frac{K'\{a\}}{K\left\{{\scriptstyle \mathsf{aw}\uparrow}\dfrac{a}{\mathsf{t}}\right\}} \longrightarrow \rho\frac{K'\left\{{\scriptstyle \mathsf{aw}\uparrow}\dfrac{a}{\mathsf{t}}\right\}}{K\{\mathsf{t}\}}$$

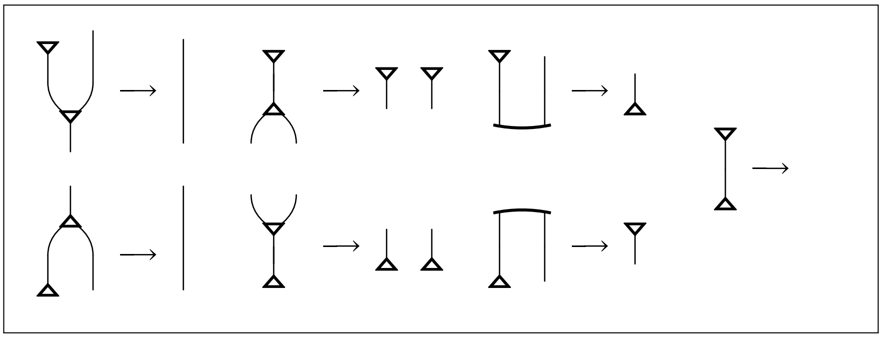


Figure 1.2: The rewriting system W in the atomic flow

We define the rewriting system $\mathsf{W} = \{\mathsf{aw}{\downarrow} - \mathsf{ac}{\downarrow}, \mathsf{aw}{\downarrow} - \mathsf{ac}{\uparrow}, \mathsf{aw}{\downarrow} - \mathsf{ai}{\uparrow}, \mathsf{aw}{\downarrow} - \mathsf{aw}{\uparrow}, \mathsf{ac}{\uparrow} - \mathsf{aw}{\uparrow}, \mathsf{ac}{\downarrow} - \mathsf{aw}{\uparrow}, \mathsf{ai}{\downarrow} - \mathsf{aw}{\uparrow}, \mathsf{aw}{\downarrow} - \rho_K, \rho_K - \mathsf{aw}{\uparrow}\}$

Proposition 1.48. *The rewriting system* W *is terminating.*

Proof. By observing the corresponding flow reductions, it is easy to see that the non-trivial reductions of W remove edges of atomic flows. Since every application of a non-trivial reduction rule reduces the number of edges of the associated flow to a derivation, and the trivial rules reduce the number of rules below weakenings and above coweakenings, termination is clear. □

Proposition 1.49. *Let* $\begin{array}{c} A \\ \phi \Vert S \\ B \end{array}$ *be a derivation, where* $\{\mathsf{aw}{\downarrow}, \mathsf{aw}{\uparrow}\} \subseteq S \subseteq \mathsf{SKS}$. *Then we can find a derivation of the following form:*

$$\begin{array}{c} A \\ \phi' \Vert S \cup \{\mathsf{s}\} \\ B \end{array} \equiv \begin{array}{c} A \\ \Vert \{\mathsf{aw}{\uparrow}\} \\ A' \\ \Vert (S \setminus \{\mathsf{aw}{\downarrow}, \mathsf{aw}{\uparrow}\}) \cup \{\mathsf{s}\} \\ B' \\ \Vert \{\mathsf{aw}{\downarrow}\} \\ B \end{array}$$

Proof. It is clear that proofs of the above form are exactly those proofs in normal form w.r.t. W. We apply W to ϕ, and, since it is terminating, we will eventually reach ϕ' of the required form. Note that the $\mathsf{aw}{\downarrow} - \mathsf{aw}{\uparrow}$ rewrite rule can introduce switches into derivations where they were not before present. □

Remark 1.50. By adding the inference rule $u\,\frac{\mathsf{f}}{\mathsf{t}}$ to a proof system (it is derivable for $\{\mathsf{s}\}$ or $\{\mathsf{m}\}$), we can guarantee that if $\phi \rightarrow^*_{\mathsf{W}} \psi$, then $|\psi| \leq |\phi|$.

Proposition 1.51. $\mathsf{aw}{\uparrow}$ *is admissible for all* $\{\mathsf{s}\} \subseteq S \subseteq \mathsf{SKS}$.

Proof. Immediate from applying Proposition 1.49 to a proof, since t cannot be the premise of an $\mathsf{aw}{\uparrow}$ instance. □

Admissibility of coweakening gives us a nice way to prove a useful proposition about KS.

Proposition 1.52. *Given a proof* $\begin{array}{c}\phi \Vert \mathsf{KS} \\ K\{A \wedge B\}\end{array}$*, we can construct two proofs,* $\begin{array}{c}\phi_l \Vert \mathsf{KS} \\ K\{A\}\end{array}$ *and* $\begin{array}{c}\phi_r \Vert \mathsf{KS} \\ K\{B\}\end{array}$*. We can then construct the following proof of $A \wedge B$:*

$$\begin{array}{c} \begin{array}{c}\phi_l \Vert \mathsf{KS} \\ K\{A\}\end{array} \wedge \begin{array}{c}\phi_r \Vert \mathsf{KS} \\ K\{B\}\end{array} \\ \Vert \{\mathsf{s},\mathsf{ac}\downarrow,\mathsf{m},\mathsf{aw}\uparrow\} \\ K\{A \wedge B\} \end{array}$$

Proof. Take

$$\begin{array}{c}\phi_l' \Vert \mathsf{KS} \\ K\left\{ {\scriptstyle =}\dfrac{A \wedge \mathsf{w}\uparrow \dfrac{B}{\mathsf{t}}}{A} \right\}\end{array} \quad \text{and} \quad \begin{array}{c}\phi_r' \Vert \mathsf{KS} \\ K\left\{ {\scriptstyle =}\dfrac{\mathsf{w}\uparrow \dfrac{A}{\mathsf{t}} \wedge B}{B} \right\}\end{array}$$

Reduce the coweakenings to atomic form, and then eliminate them from each proof with Proposition 1.51.

The second part is easy to show by induction on $K\{\ \}$. There is nothing to do for the base case and the two inductive steps are given below.

$$ {\scriptstyle =}\dfrac{(K'\{A\} \wedge C) \wedge (K'\{B\} \wedge C)}{\begin{array}{c}K'\{A\} \wedge K'\{B\} \\ {\scriptstyle IH}\Vert \\ K'\{A \wedge B\}\end{array} \wedge \left(\begin{array}{c} C \wedge C \\ \Vert \{\mathsf{aw}\uparrow\} \\ C \end{array}\right)} \quad \text{and} \quad {\scriptstyle \mathsf{s}^2}\dfrac{(K'\{A\} \vee C) \wedge (K'\{B\} \vee C)}{\begin{array}{c}K'\{A\} \wedge K'\{B\} \\ {\scriptstyle IH}\Vert \\ K'\{A \wedge B\}\end{array} \vee \left(\begin{array}{c} C \vee C \\ \Vert \{\mathsf{ac}\downarrow,\mathsf{m}\} \\ C \end{array}\right)} $$

□

1.4 The Experiments Method for Cut Elimination

1.4.1 Elimination of the 'up'-rules.

For both propositional and predicate logic cut elimination is fairly straightforward: one can achieve it by translation into the sequent calculus, using Gentzen's *Hauptsatz* [Gen64] and translating back into a deep inference system [Brü03a]. However, as well as this method being inelegant and uninteresting, it has been shown that there are cut elimination methods for SKS that can be done in quasipolynomial time [Bru+16; Jeř08], whereas all known cut elimination methods for shallow inference systems require exponential time. The experiments method we describe below does not give these complexity advantages, but, beyond a certain elegance, achieves a sort of confluence for cut elimination, given certain provisos. Although this method of cut elimination has not been published before, it has been informally described by Guglielmi and others in talks [Gug+10; RG15].

1.4.2 Experiments

Before proceeding to the elimination of $\mathsf{ai}\!\uparrow$, we need a few preliminary lemmas.

Lemma 1.53. *For any formula context $K\{\ \}$ and any formula A there are derivations*

$$\begin{array}{c} K\{\mathsf{t}\} \wedge A \\ \| \{\mathsf{s}\} \\ K\{A\} \end{array} \quad \textit{and} \quad \begin{array}{c} K\{A\} \\ \| \{\mathsf{s}\} \\ K\{\mathsf{f}\} \vee A \end{array}$$

.

Proof. We show that we can construct $\begin{array}{c} K\{\mathsf{t}\} \wedge A \\ \| \{\mathsf{s}\} \\ K\{A\} \end{array}$ by induction on the size of $K\{\ \}$.

If $K\{\ \} = \{\ \}$ then the derivations are simply $= \dfrac{\mathsf{t} \wedge A}{A}$ and $= \dfrac{A}{\mathsf{f} \vee A}$.

The two inductive steps are as follows:

$$\mathsf{s}\,\dfrac{\dfrac{K\{\mathsf{t}\}}{B \vee K'\{\mathsf{t}\}} \wedge A}{\dfrac{B \vee \boxed{\begin{array}{c} K'\{\mathsf{t}\} \wedge A \\ IH \| \{\mathsf{s}\} \\ K'\{A\} \end{array}}}{K\{A\}}} \quad \text{and} \quad =\dfrac{\dfrac{K\{\mathsf{t}\}}{B \vee K'\{\mathsf{t}\}} \wedge A}{\dfrac{B \wedge \boxed{\begin{array}{c} K'\{\mathsf{t}\} \wedge A \\ IH \| \{\mathsf{s}\} \\ K'\{A\} \end{array}}}{K\{A\}}}$$

The proof for $\begin{array}{c} K\{A\} \\ \| \{\mathsf{s}\} \\ K\{\mathsf{f}\} \wedge A \end{array}$ is dual. $\square$

Lemma 1.54. *Given an* SKS *(or any system* $\{\mathsf{s}\} \subseteq \mathsf{S} \subseteq \mathsf{SKS}$*) derivation* $\begin{array}{c} A \\ \phi \| \mathsf{SKS} \\ B \end{array}$*, we can construct derivations of the following form:*

$$\begin{array}{c} = \dfrac{A}{A \wedge \boxed{\mathsf{ai}\!\downarrow \dfrac{\mathsf{t}}{a_1 \vee \bar{a}_1}} \wedge \ldots \wedge \boxed{\mathsf{ai}\!\downarrow \dfrac{\mathsf{t}}{a_n \vee \bar{a}_n}}} \\ \phi' \Big\| \mathsf{SKS} \backslash \{\mathsf{ai}\!\downarrow, \mathsf{ai}\!\uparrow\} \\ = \dfrac{B \vee \boxed{\mathsf{ai}\!\uparrow \dfrac{b_1 \wedge \bar{b}_1}{\mathsf{f}}} \vee \ldots \vee \boxed{\mathsf{ai}\!\uparrow \dfrac{b_n \wedge \bar{b}_n}{\mathsf{f}}}}{B} \end{array}$$

Proof. We will just show that $\mathsf{ai}\!\downarrow$ rules can be pushed up: $\mathsf{ai}\!\uparrow$ rules can be pushed down in exactly the same way, dually. We proceed by induction on the number of ai↓ instances in ϕ. The inductive step is as follows:

$$\phi \longrightarrow^*_{\mathsf{Seq}} \begin{array}{c} A \\ \phi_1 \,\Vert\, \mathsf{SKS} \\ K\left\{ \mathsf{ai}\!\downarrow \dfrac{\mathsf{t}}{a_n \vee \bar{a}_n} \right\} \\ \phi_2 \,\Vert\, \mathsf{SKS}\setminus\{\mathsf{ai}\!\downarrow\} \\ B \end{array} \xrightarrow{\text{Lem 1.53}} \begin{array}{c} = \dfrac{A}{\boxed{\begin{array}{c} A \\ \phi_1 \,\Vert\, \mathsf{SKS} \\ K\{\mathsf{t}\} \end{array}} \wedge \boxed{\mathsf{ai}\!\downarrow \dfrac{\mathsf{t}}{a_n \vee \bar{a}_n}}} \\ \Vert\, \{\mathsf{s}\} \\ K\{a_n \vee \bar{a}_n\} \\ \phi_2 \,\Vert \\ B \end{array} \xrightarrow{IH}$$

$$\begin{array}{c} \boxed{= \dfrac{A}{\boxed{\begin{array}{c} \dfrac{A}{A \wedge \boxed{\mathsf{ai}\!\downarrow \dfrac{\mathsf{t}}{a_1 \vee \bar{a}_1}} \wedge \ldots \wedge \boxed{\mathsf{ai}\!\downarrow \dfrac{\mathsf{t}}{a_{n-1} \vee \bar{a}_{n-1}}}} \\ \phi_1' \,\Vert\, \mathsf{SKS}\setminus\{\mathsf{ai}\!\downarrow\} \\ K\{\mathsf{t}\} \end{array}} \wedge \boxed{\mathsf{ai}\!\downarrow \dfrac{\mathsf{t}}{a_n \vee \bar{a}_n}}}} \\ \Vert\, \{\mathsf{s}\} \\ K\{a_n \vee \bar{a}_n\} \\ \phi_2 \,\Vert \\ B \end{array}$$

$$\longrightarrow \begin{array}{c} \dfrac{A}{A \wedge \boxed{\mathsf{ai}\!\downarrow \dfrac{\mathsf{t}}{a_1 \vee \bar{a}_1}} \wedge \ldots \wedge \boxed{\mathsf{ai}\!\downarrow \dfrac{\mathsf{t}}{a_n \vee \bar{a}_n}}} \\ \phi' \,\Vert\, \mathsf{SKS}\setminus\{\mathsf{ai}\!\downarrow\} \\ B \end{array}$$

□

We are now ready to prove our cut elimination result with the Experiments method. The general idea is that for each cut, we create two versions of the proof, the intuition being that for each cut $\mathsf{ai}\!\uparrow \dfrac{a \wedge \bar{a}}{\mathsf{f}}$, either a or $\bar{a}$ is 'true' (in the Tarskian sense), and each version corresponds to one of these possibilities. We then use an identity and a contraction to disjunct the two proofs, creating a proof with one fewer cut instance.

Theorem 1.55 (Cut Elimination with the Experiments method)**.** ai↑ *is admissible for all systems* $\mathsf{KS} \subseteq \mathsf{S} \subseteq \mathsf{SKS}\setminus\{\mathsf{ai}\!\uparrow\}$

Proof. We proceed by induction on the number of distinct atoms in ϕ for which there is at least one instance of ai↓.

1. Let $\begin{array}{c} \phi \,\Vert\, \mathsf{S} \\ A \end{array}$. First, we use Lemma 1.54 to push the ai↓ rules to the top of the

proof, to give us:

$$\begin{array}{c} \overline{\|}\{\mathsf{ai}\downarrow\} \\ (a_i \vee \bar{a}_i)^n \\ \psi \| \mathsf{S}\setminus\{\mathsf{ai}\downarrow\} \\ A \end{array}$$

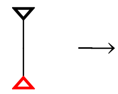
where $(a_i \vee \bar{a}_i)^n \equiv ((a_1 \vee \bar{a}_1) \wedge \ldots \wedge (a_n \vee \bar{a}_n))$ with the a_i not necessarily distinct.

2. Pick an arbitrary atom a in ψ s.t. there is at least one instance of $\mathsf{ai}\uparrow$ involving a. Replace every such instance of $\mathsf{ai}\uparrow$ like so:

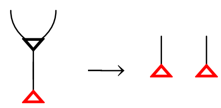
$$= \frac{\mathsf{f}\,\dfrac{a}{\mathsf{f}} \wedge \mathsf{aw}\uparrow\dfrac{\bar{a}}{\mathsf{t}}}{\mathsf{f}}$$

Note that we have an unsound rule, f. We can use the following rewrites to push it to the top of ψ (along with similar trivial reductions to W):

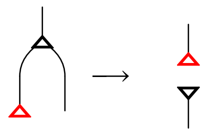
$\mathsf{aw}\downarrow - \mathsf{f}$: $\quad \mathsf{aw}\downarrow\dfrac{\mathsf{f}}{\mathsf{f}\,\dfrac{a}{\mathsf{f}}} \longrightarrow \mathsf{f}$



$\mathsf{ac}\downarrow - \mathsf{f}$: $\quad \mathsf{ac}\downarrow\dfrac{a \vee a}{\mathsf{f}\,\dfrac{a}{\mathsf{f}}} \longrightarrow \mathsf{f}\,\dfrac{a}{\mathsf{f}} \vee \mathsf{f}\,\dfrac{a}{\mathsf{f}}$

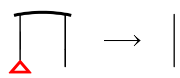


$\mathsf{ac}\downarrow - \mathsf{f}$: $\quad \mathsf{ac}\uparrow\dfrac{a}{\mathsf{f}\,\dfrac{a}{\mathsf{f}} \wedge a} \longrightarrow = \dfrac{\mathsf{f}\,\dfrac{a}{\mathsf{f}}}{\mathsf{f} \wedge \mathsf{aw}\downarrow\dfrac{\mathsf{f}}{a}}$



If f reaches the top of ψ, it must hit an $\mathsf{ai}\downarrow$ rule. We perform the following rewrite each time this occurs:

$\mathsf{ai}\downarrow - \mathsf{f}$: $\quad \mathsf{ai}\downarrow\dfrac{\mathsf{t}}{\mathsf{f}\,\dfrac{a}{\mathsf{f}} \vee \bar{a}} \longrightarrow = \dfrac{\bar{a}}{\mathsf{f} \vee \bar{a}}$



We now have a derivation of the following form (if necessary, $\mathsf{aw}\uparrow$ can be removed by Proposition 1.51):

$$\begin{array}{c} (\bar{a})^j \\ \psi^- \| \mathsf{S} \\ A \end{array} \equiv \begin{array}{c} (\bar{a})^j \\ \| \{\mathsf{ai}\downarrow\} \\ (a_i \vee \bar{a}_i)^{n-j} \\ \psi' \| \mathsf{S}\setminus\{\mathsf{ai}\downarrow\} \\ A \end{array}$$

Similarly, we can construct $\begin{array}{c} (a)^j \\ \psi^+ \| \mathsf{S} \\ A \end{array}$

3. We now combine ψ^+ and ψ^- together to make a new proof:

$$\phi' \equiv \begin{array}{c} \Big\Vert\{\mathsf{ai}\downarrow,\mathsf{ac}\downarrow\} \\ \boxed{\begin{array}{c} (\bar{a})^j \\ \psi^- \Vert \mathsf{S} \\ A \end{array} \vee \begin{array}{c} (a)^j \\ \psi^+ \Vert \mathsf{S} \\ A \end{array}} \\ \Vert\{\mathsf{ac}\downarrow,\mathsf{m}\} \\ A \end{array}$$

As ϕ' contains cuts on all the atoms in ϕ except a, we are done.

□

Proof of Theorem 1.45 (Cut Elimination for CPL). Immediate from Theorem 1.55 and Proposition 1.46. □

We can also use the experiments without an inductive argument, to get a cut-free proof that is in some sense canonical. Instead of working on one atom at a time we consider all atoms involved in cuts at once: $(a_1, \ldots, a_n)$. Then, for each of the 2^n truth assignments $A_i : \mathcal{A} \to \mathcal{U}$, we construct a cut-free derivation

$$\begin{array}{c} \left(b_j\right)^{m_{(i,1)}} \wedge \left(\bar{b}_k\right)^{m_{(i,2)}} \\ \phi_i \Vert \mathsf{S}\backslash\{\mathsf{ai}\uparrow\} \\ A \end{array}$$

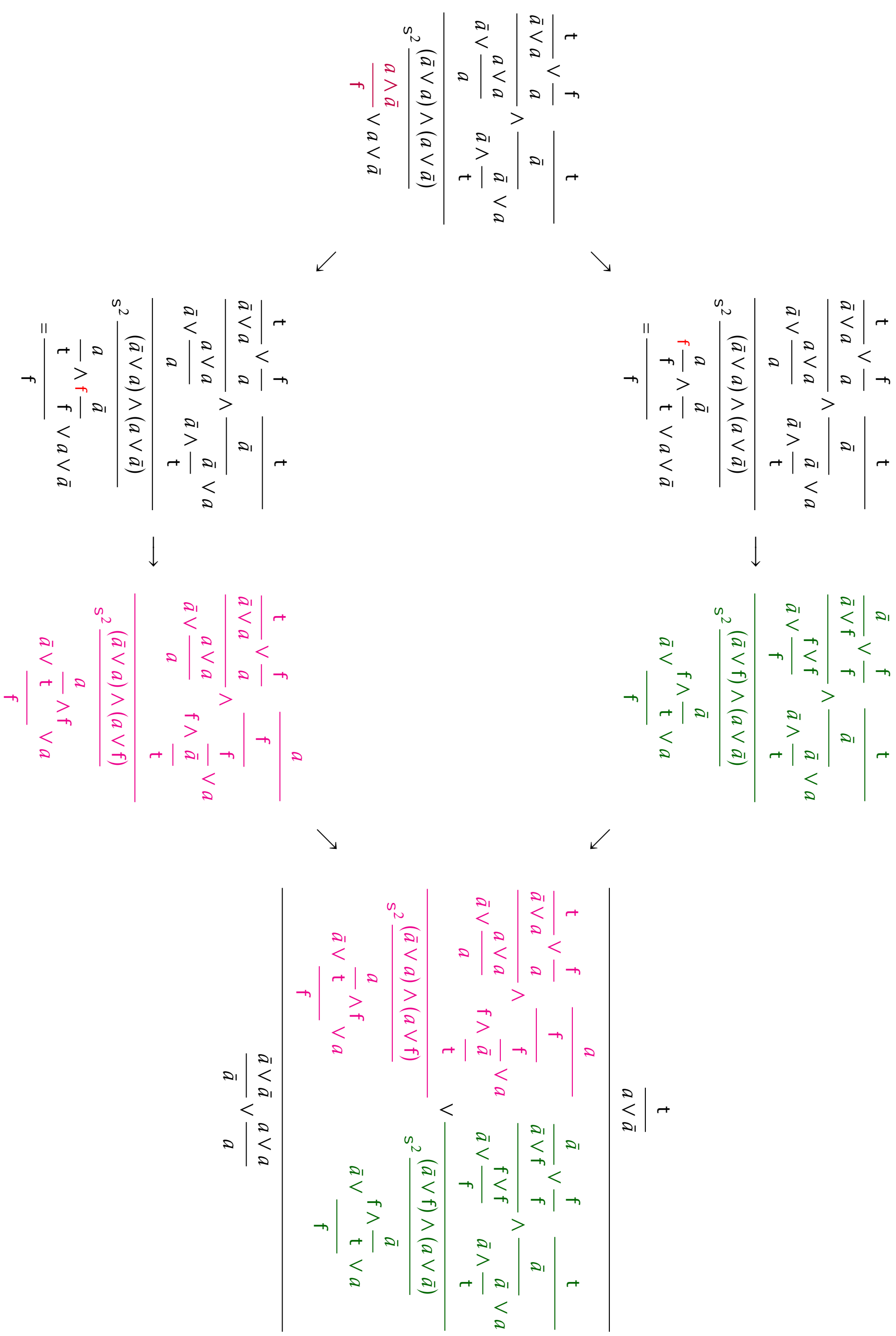
where $b_j \in \{a_j \mid A_i(a_j) = \mathsf{t}\}$ and $b_k \in \{a_k \mid A_i(a_k) = \mathsf{f}\}$. We can then construct a cut-free proof of A:

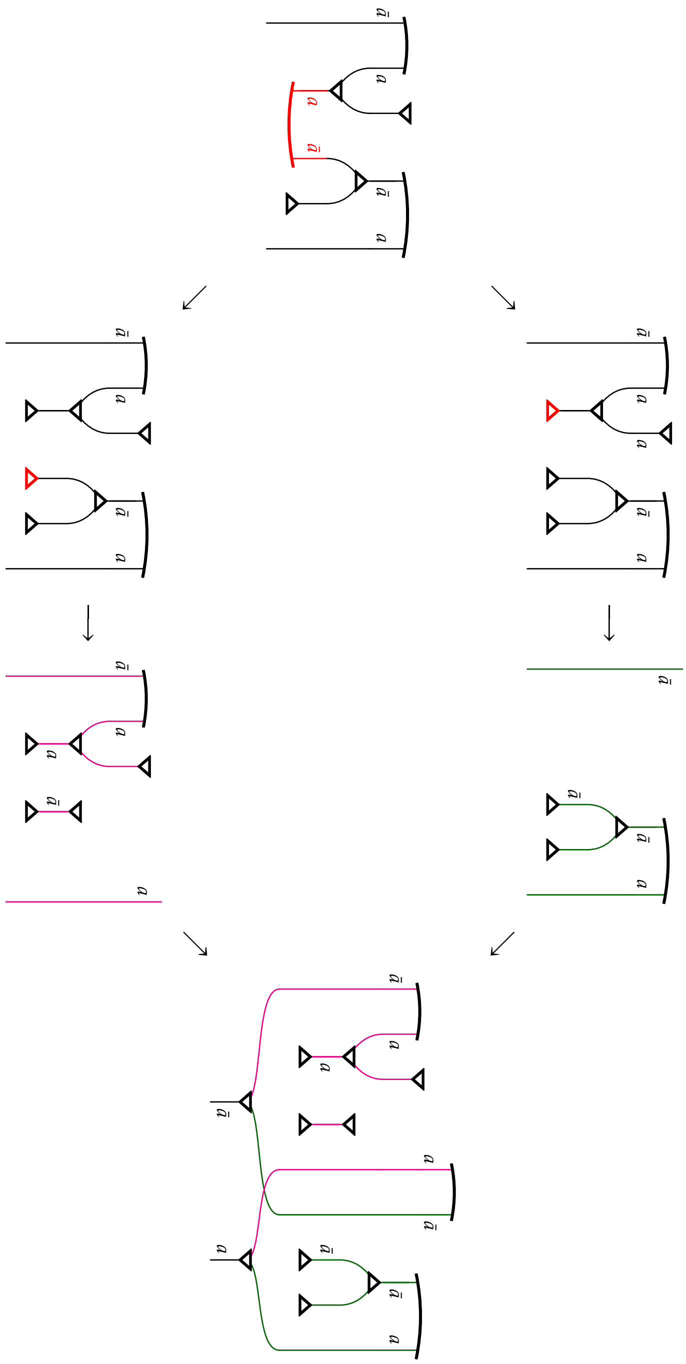
$$\begin{array}{c} TT_\phi \Big\Vert \mathsf{KS} \\ \boxed{\begin{array}{c} \left(a_j\right)^{m_{(1,1)}} \wedge (\bar{a}_k)^{m_{(1,2)}} \\ \phi_1 \Vert \mathsf{S}\backslash\{\mathsf{ai}\uparrow\} \\ A \end{array} \vee \ldots \vee \begin{array}{c} \left(a_j\right)^{m_{(2^n,1)}} \wedge (\bar{a}_k)^{m_{(2^n,2)}} \\ \phi_{2^n} \Vert \mathsf{S}\backslash\{\mathsf{ai}\uparrow\} \\ A \end{array}} \\ \Vert\{\mathsf{ac}\downarrow,\mathsf{m}\} \\ A \end{array}$$

where TT_ϕ is a proof of the relevant truth-table tautology, as in [Das14b, Lemma 8.2].

It is simple to check that the f reduction system is confluent, so the cut-free proof is essentially canonical up to commutativity of the ϕ_i.

Example 1.56. Below we show the experiments method on a proof with one atomic cut.





Chapter 2

Decomposition via Cycle Removal

In the previous chapter, we showed how cuts can be reduced to atomic form and then eliminated from an open deduction proof. In this chapter, we will not focus on cut elimination itself, but on *decomposition*: pushing contractions to the bottom of a proof.

To do so, we define a rewriting system of atomic contractions, whose normal forms are precisely decomposed proofs. However, unlike the weakening rewriting system W we saw in the last chapter, this rewriting system is not even weakly normalising. We isolate the source of infinite reduction sequences—a certain sort of cycle in proofs—and show how we can obtain a cycle-free proof on which the contraction rewriting system is strongly normalizing.

2.1 Decomposition and Cycles

Much of the material in the following section originates in [GG08], reworked in numerous subsequent papers and again here.

First, we will define the rewriting system we will use to decompose proofs.

Definition 2.1. We define the following reduction rules for SKS:

$$\mathsf{ac}{\downarrow} - \mathsf{ac}{\uparrow} : \quad \mathsf{ac}{\uparrow}\frac{\mathsf{ac}{\downarrow}\dfrac{a \vee a}{a}}{a \wedge a} \quad \longrightarrow \quad \mathsf{m}\frac{\mathsf{ac}{\uparrow}\dfrac{a}{a \wedge a} \vee \mathsf{ac}{\uparrow}\dfrac{a}{a \wedge a}}{\mathsf{ac}{\downarrow}\dfrac{a \vee a}{a} \wedge \mathsf{ac}{\downarrow}\dfrac{a \vee a}{a}}$$

$$\mathsf{ac}{\downarrow} - \mathsf{ai}{\uparrow} : \quad \mathsf{ai}{\uparrow}\frac{\mathsf{ac}{\downarrow}\dfrac{a \vee a}{a} \wedge \bar{a}}{\mathsf{f}} \quad \longrightarrow \quad = \frac{2\mathsf{s}\dfrac{(a \vee a) \vee \mathsf{ac}{\uparrow}\dfrac{\bar{a}}{\bar{a} \wedge \bar{a}}}{\mathsf{ai}{\uparrow}\dfrac{a \wedge \bar{a}}{\mathsf{f}} \vee \mathsf{ai}{\uparrow}\dfrac{a \wedge \bar{a}}{\mathsf{f}}}}{\mathsf{f}}$$

$$\mathsf{ac}{\downarrow} - \mathsf{aw}{\uparrow} : \quad \mathsf{aw}{\uparrow}\frac{\mathsf{ac}{\downarrow}\dfrac{a \vee a}{a}}{\mathsf{t}} \quad \longrightarrow \quad = \frac{\mathsf{aw}{\uparrow}\dfrac{a}{\mathsf{t}} \vee \mathsf{aw}{\uparrow}\dfrac{a}{\mathsf{t}}}{\mathsf{t}}$$

And their duals:

$$\mathsf{ai}{\downarrow} - \mathsf{ac}{\uparrow} : \quad \mathsf{ai}{\downarrow}\frac{\mathsf{t}}{\mathsf{ac}{\uparrow}\dfrac{a}{a \wedge a} \vee \bar{a}} \quad \longrightarrow \quad 2\mathsf{s}\frac{= \dfrac{\mathsf{t}}{\mathsf{ai}{\downarrow}\dfrac{\mathsf{t}}{a \vee \bar{a}} \wedge \mathsf{ai}{\downarrow}\dfrac{\mathsf{t}}{a \vee \bar{a}}}}{(a \wedge a) \wedge \mathsf{ac}{\downarrow}\dfrac{\bar{a} \vee \bar{a}}{\bar{a}}}$$

$$\mathsf{aw}{\downarrow} - \mathsf{ac}{\uparrow} : \quad \mathsf{ac}{\uparrow}\frac{\mathsf{aw}{\downarrow}\dfrac{\mathsf{f}}{a}}{a \wedge a} \quad \longrightarrow \quad = \frac{\mathsf{f}}{\mathsf{aw}{\downarrow}\dfrac{\mathsf{f}}{a} \wedge \mathsf{aw}{\downarrow}\dfrac{\mathsf{f}}{a}}$$

Last, we define the trivial family of reduction rules:

$$\mathsf{ac}{\downarrow} - \rho_H : \quad \rho\frac{H\left\{\mathsf{ac}{\downarrow}\dfrac{a \vee a}{a}\right\}}{H'\{a\}} \quad \longrightarrow \quad \rho\frac{H\{a \vee a\}}{H'\left\{\mathsf{ac}{\downarrow}\dfrac{a \vee a}{a}\right\}}$$

$$\rho_H - \mathsf{ac}{\uparrow} : \quad \rho\frac{H'\{a\}}{H\left\{\mathsf{ac}{\uparrow}\dfrac{a}{a \wedge a}\right\}} \quad \longrightarrow \quad \rho\frac{H'\left\{\mathsf{ac}{\uparrow}\dfrac{a}{a \wedge a}\right\}}{H\{a \wedge a\}}$$

These simply correspond to drawing the (co)contraction node lower (higher) in the atomic flow.

Definition 2.2. We define *rewriting system* C for SKS as the rewriting system given by the reduction rules of Definition 2.1.

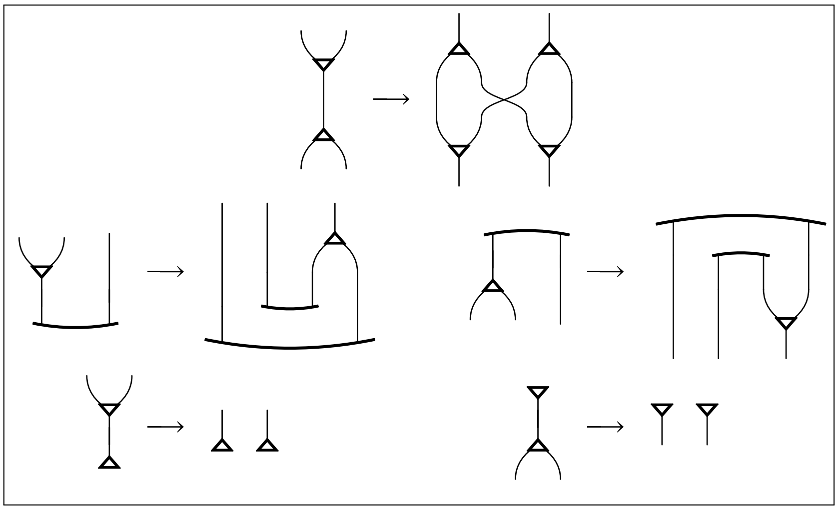


Figure 2.1: The rewriting system C in the atomic flow.

Observation 2.3. Applying C to a proof may increase the size exponentially, through the crossings of contractions and cocontractions as shown in Figure 2.2. We call the contraction-cocontraction pairs on the left *sausages*.

Definition 2.4. Let ϕ be a SKS derivation in the following form:

$$\begin{array}{c} A \\ \| \{\mathsf{ac}\uparrow\} \\ A' \\ \| \mathsf{SKS} \setminus \{\mathsf{ac}\downarrow, \mathsf{ac}\uparrow\} \\ B' \\ \| \{\mathsf{ac}\downarrow\} \\ B \end{array}$$

We say that ϕ is *decomposed*.

Proposition 2.5. *A derivation ϕ is decomposed iff it is in normal form w.r.t.* C.

Definition 2.6. Let ϕ be a SKS derivation in the following form:

$$\begin{array}{c} A \\ \| \{\mathsf{aw}\uparrow\} \\ A' \\ \| \{\mathsf{ac}\uparrow\} \\ A'' \\ \| \mathsf{SKS} \setminus \{\mathsf{ac}\downarrow, \mathsf{ac}\uparrow, \mathsf{aw}\downarrow, \mathsf{aw}\uparrow\} \\ B'' \\ \| \{\mathsf{ac}\downarrow\} \\ B \\ \| \{\mathsf{aw}\downarrow\} \\ B \end{array}$$

We say that ϕ is *strongly decomposed*.

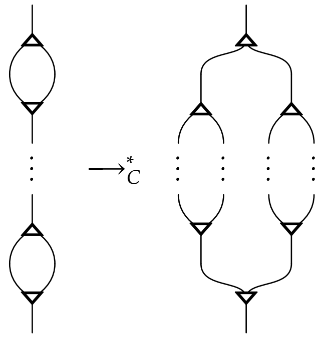


Figure 2.2: Exponential blow-up caused by sausages.

Proposition 2.7. *A derivation ϕ is strongly decomposed iff it is in normal form w.r.t.* C *and* W.

2.1.1 ai-cycles

Unlike with W, C does not obviously terminate. In fact, it is not difficult to find infinite reduction sequences for C: these turn out to be exactly those proofs that contain a particular kind of *cycle* in their atomic flow. For precision, we need to develop a little more theory for atomic flows.

Definition 2.8. Given an edge ϵ in an atomic flow, we define $\mathsf{up}(\epsilon)$ as the upper vertex it is connected to, and $\mathsf{lo}(\epsilon)$ as the lower vertex it is connected to.

Definition 2.9. Given a sequence of distinct edges $\epsilon_1, \ldots, \epsilon_n$ such that $\mathsf{lo}(\epsilon_i) = \mathsf{up}(\epsilon_{i+1})$ for $1 \leq i < n$, we say that $\epsilon_1, \ldots, \epsilon_n$ is a *straight path of length n from* $\mathsf{up}(\epsilon_1)$ *to* $\mathsf{lo}(\epsilon_n)$, and that $\epsilon_n, \ldots, \epsilon_1$ is a *straight path of length n from* $\mathsf{lo}(\epsilon_n)$ *to* $\mathsf{up}(\epsilon_1)$.

Given a sequence of edges $\epsilon_1, \ldots, \epsilon_n$, we say that $\epsilon_1, \ldots, \epsilon_n$ is a *path of length n from vertex v_1 to vertex v_2* if it is a straight path from v_1 to v_2 or if there exists a vertex v labelled by $\mathsf{ai}{\uparrow}$ or $\mathsf{ai}{\downarrow}$ such that $\epsilon_1, \ldots, \epsilon_h$ is a straight path from v_1 to v, $\epsilon_h \neq \epsilon_{h+1}$, and $\epsilon_{h+1}, \ldots, \epsilon_n$ is a path from v to v_2.

Definition 2.10. A path from v to v is called an ai-*cycle*.

A path of length n is *maximal* if no path containing its edges as a subsequence has length greater than n. A path of length n from v is *maximal from v* if no path from v containing its edges as a subsequence has length greater than n.

We can now define the measure used to prove termination of C.

Definition 2.11. Let v be a vertex labelled with $\mathsf{ac}{\downarrow}$ in a flow, with no path from v we define its *rank* as the sum of the lengths of the maximal from v paths $\epsilon_1, \ldots, \epsilon_n$ such that $\mathsf{up}(\epsilon_1) = v$.

Dually, given a vertex v labelled with $\mathsf{ac}{\uparrow}$ in a flow, we define its *rank* as the sum of the lengths of the maximal from v paths $\epsilon_1, \ldots, \epsilon_n$ such that $\mathsf{lo}(\epsilon_1) = v$.

Remark 2.12. Clearly, a vertex that is part of a ai-cycle has "infinite rank", in a sense. But, since paths are defined to be finite, we do not formally define the notion of infinite rank. Therefore the definition of rank only makes sense for cycle-free flows.

Definition 2.13. We say that a derivation contains an ai-cycle if its atomic flow contains an ai-cycle.

2.1.2 Termination of C

We can easily see that if we repeatedly perform rewrites exclusively on a contraction inside an ai-cycle, such as in Figure 2.3, then the rewriting procedure will not terminate.

In the absence of such cycles however, the rewriting always terminates. We give a sketch of the proof, a more precise version can be found in [GG08] and elsewhere.

Theorem 2.14. *The rewriting system* C *is terminating on the set of* ai*-cycle-free derivations.*

Proof. (Sketch) First, we observe that it is clear by inspection of the reduction rules that the rank of (co)contractions not involved in the reduction stays the same.

Given an ai-cycle-free derivation ϕ, we consider the lexicographic order on (r,d). r is the sum of the ranks of the contractions and cocontractions in ϕ, and d is the sum of the number of rules below each contraction and the number of rules above each cocontraction when sequentialising ϕ.

We describe how each application of a reduction of C reduces (r,d):

- Applications of the rules $\mathsf{ac}{\downarrow}-\mathsf{ac}{\uparrow}$, $\mathsf{ac}{\downarrow}-\mathsf{ai}{\downarrow}$ and $\mathsf{ai}{\downarrow}-\mathsf{ac}{\uparrow}$ reduce r in the absence of ai-cycles as is shown in the proof of Theorem 7.2.3 of [Gun09].
- Applications of the rules $\mathsf{ac}{\downarrow}-\mathsf{aw}{\uparrow}$ and $\mathsf{aw}{\downarrow}-\mathsf{ac}{\uparrow}$ reduce r since they remove contractions and cocontractions.
- Applications of the rules $\mathsf{ac}{\downarrow}-\rho_H$ and $\rho_H-\mathsf{ac}{\uparrow}$ trivially maintain r and reduce d.

□

Evidently, ai-cycles are removed through cut-elimination, since they are caused by the connection of a cut and an introduction. In this paper we will present a local procedure to remove cycles that does not involve cut-elimination, thus proving the independence of decomposition from cut-elimination.

To improve this decomposition result, it can also be shown that (co)weakenings can be permuted to the bottom (top) of a derivation using rewriting system W above [Das15].

Note that the reductions of system W do not introduce atomic (co)contractions or medials. Thus, we get the following theorem.

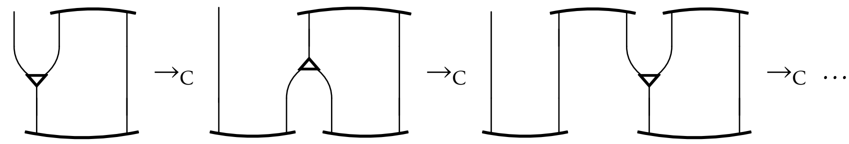


Figure 2.3: A flow that does not terminate under C.

Theorem 2.15. *Given an* SKS *derivation ϕ from A to B not containing* ai*-cycles, we can obtain a strongly decomposed derivation ϕ from A to B.*

Proof. By applying system C followed by system W to ϕ. ☐

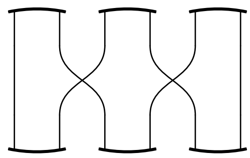
2.1.3 n-Multicycles and Unicycles

To ease proofs later, we show how to reduce complex cycles to those with a single ai↓ rule and a single ai↑ rule, where the ai↑ rule is not shared by any other cycles.

Definition 2.16. We say that an ai-cycle is a *unicycle* when it only crosses a single ai↓ node and a single ai↑ node. We say that an ai-cycle is an *n-multicycle* when it crosses n distinct ai↓ nodes and n distinct ai↑ nodes.

Example 2.17. Each flow in Figure 2.3 contains one unicycle.

The cycle below is a 3-multicycle.



Whereas in unicycles it is clear that the the logical relation between the two atom occurrences involved in the cycle must change from a conjunction $\vee$ to a disjunction $\wedge$, in multicycles it is not necessarily so. It is however easy to transform multicycles into unicycles using standard SKS derivation transformations.

Lemma 2.18. *Given a derivation ϕ with an n-multicycle ($n > 1$), there exists a derivation ψ where the multicycle is replaced by n unicycles. We do so by collapsing multiple identities on one atom into one.*

Proof. We show that a derivation with an n-multicycle can be converted into a derivation with an n_1-multicycle and a n_2 unicycle, with $n_1 + n_2 = n$. Clearly, the lemma follows from this. Let $\begin{array}{c} A \\ \phi \Vert \\ B \end{array}$ be a derivation with an n-multicycle. We denote by ${\color{red}a}, {\color{red}\bar{a}}, {\color{blue}a}, {\color{blue}\bar{a}}$ positive and negative instances of the same atom, in the same cycle, but use the colours to distinguish different instances of the identity rule.

By Lemma 1.54, we can construct:

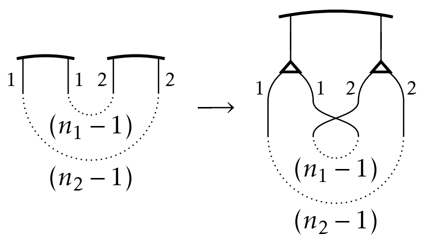
$$\begin{array}{c} A \wedge \boxed{\mathsf{ai}{\downarrow}\,\dfrac{\mathsf{t}}{{\color{red}a} \vee {\color{red}\bar{a}}}} \wedge \boxed{\mathsf{ai}{\downarrow}\,\dfrac{\mathsf{t}}{{\color{blue}a} \vee {\color{blue}\bar{a}}}} \\ \phi' \Vert \\ B \end{array} .$$

Thus, we take ψ to be:

$$\psi \equiv \begin{array}{c} A \wedge \boxed{\mathsf{m}\,\dfrac{\boxed{\mathsf{ai}{\downarrow}\,\dfrac{\mathsf{t}}{\boxed{\mathsf{ac}{\downarrow}\,\dfrac{a}{{\color{red}a} \wedge {\color{blue}a}}} \vee \boxed{\mathsf{ac}{\downarrow}\,\dfrac{\bar{a}}{{\color{red}\bar{a}} \wedge {\color{blue}\bar{a}}}}}}}{({\color{red}a} \vee {\color{red}\bar{a}}) \wedge ({\color{blue}a} \vee {\color{blue}\bar{a}})}} \\ \phi' \Vert \\ B \end{array} .$$

□

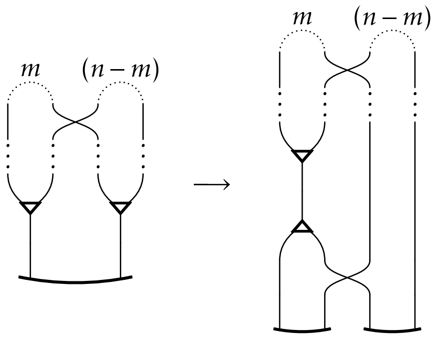
In this way, we transform an n-multicycle into an n_1-multicycle and an n_2-multicycle that share an identity:



Lemma 2.19. *Given a derivation ϕ with n unicycles with different identities that share a cut, there exists a derivation ψ with the same premise and conclusion where they are replaced by n unicycles that do not share a cut.*

Proof. We will show that if we have a derivation with one instance of n unicycles sharing a cut, we can construct a derivation with at most $n-1$ unicycles sharing a cut. The lemma clearly follows from this.

For there to be more than one unicycle sharing a cut, each edge of the cut must contract at least once. If there are cocontractions between both cuts and contractions, we use the $\mathsf{ac}{\downarrow} - \mathsf{ac}{\uparrow}$ rewrite to permute the contraction down to the cut—since this does not change any paths, it cannot affect the cycles. Then, once the contraction is directly above the cut, we apply the $\mathsf{ac}{\downarrow} - \mathsf{ai}{\uparrow}$ rewrite once:



□

Proposition 2.20. *Let ϕ be an* SKS *derivation. We can construct a derivation ϕ' with the same premise and conclusion such that:*

1. *ϕ' contains no n-multicycles for $n \geq 2$.*
2. *No atomic cut in ϕ is a node in more than one cycle.*

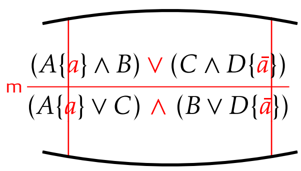
Proof. Immediate from Lemmas 2.18 and 2.19. □

Our strategy for removing cycles is based on a simple observation: that at the top of the cycle, the positive and negative occurrence of the atom are introduced with an $\vee$ between them, whereas at the bottom of the cycle the occurrences are destroyed with a $\wedge$. Therefore, in a unicycle, there must be at least one point in the derivation where the relation between a and $\bar{a}$ changes from $\vee$ to $\wedge$. A cursory inspection of SKS shows that the medial rule is the only inference rule that could possibly cause this change. In fact, the medial rule has been shown to be, in a certain sense, the canonical inference rule with this behaviour [DS16].

Definition 2.21. Let ϕ be a derivation containing a unicycle, represented by the ai-cycle $\epsilon_1, \ldots, \epsilon_n$ in its atomic flow. The *critical medial* for this unicycle is the lowest instance of a rule

$$\mathsf{m}\,\frac{(A\{a\} \wedge B) \vee (C \wedge D\{\bar{a}\})}{(A\{a\} \vee C) \wedge (B \vee D\{\bar{a}\})}$$

in ϕ where the occurrences of a and $\bar{a}$ are represented in the atomic flow by one of the edges belonging to the ai-cycle $\epsilon_1, \ldots, \epsilon_n$.



Following this observation, a strategy one can take to remove cycles becomes clear: we can permute the critical instances of the medial rule downwards (or upwards) in a proof. When the corresponding cut is reached, it is 'broken' by the critical medial, and the cycle can then be removed by performing standard deep inference rewriting techniques. It is by no means obvious that this suffices to remove cycles, but we will show that it in fact does.

Together with Andrea Aler Tubella and Alessio Guglielmi, two procedures to remove critical medials, and therefore ai-cycles, from derivations have been found, effectively showing the independence of decomposition and cut elimination. The original procedure is shorter, but particular to propositional logic and not of great interest in itself, published in [ATGR17]. Here we show a second procedure that is more complex, but has greater mathematical interest and generalisability. The idea has already been presented in Aler Tubella's thesis [AT17] (and a journal paper is also in preparation [ATGR18]), but the method shown here is sufficiently different to warrant a full treatment—in particular, the concept of the merge contraction is developed in far greater detail than in previous work.

2.2 Propositional Merge Contractions

In what follows we will present a rewriting system that will allow us to permute critical medials downwards in a derivation. Since medials cannot be permuted downwards past instances of the rules $\mathsf{ac}\!\downarrow$ and certain equality rules, we will permute them in the guise of a more general inference rule that we call a *merge contraction*: they are rules that correspond to particular nestings of instances of m, $\mathsf{ac}\!\downarrow$ and = and which permute with every SKS rule.

2.2.1 Merge contractions and contractive derivations

In order to motivate merge contractions and also to greatly simplify the case analysis of the reduction rules, we will exploit a common feature of most deep inference systems: that we can provide proof systems where all rules besides the atomic ones can be expressed by a single inference rule shape [ATG18]

Definition 2.22. An inference rule has the *medial shape* if it is in the following form:

$$\frac{(A\ \alpha\ B)\ \beta\ (C\ \alpha'\ D)}{(A\ \beta\ C)\ \alpha\ (B\ \beta'\ D)}$$

where A, B, C, D are formulae and $\alpha, \beta, \alpha', \beta'$ are connectives.

Definition 2.23. We define four new rules, each having the medial shape:

$$\vee\mathsf{c}\!\downarrow\frac{(A \vee B) \vee (C \vee D)}{(A \vee C) \vee (B \vee D)} \qquad \wedge\mathsf{c}\!\uparrow\frac{(A \wedge B) \wedge (C \wedge D)}{(A \wedge C) \wedge (B \wedge D)}$$

$$\mathsf{s}\!\downarrow\frac{(A \vee B) \wedge (C \vee D)}{(A \wedge C) \vee (B \vee D)} \qquad \mathsf{s}\!\uparrow\frac{(A \vee B) \wedge (C \wedge D)}{(A \wedge C) \vee (B \wedge D)}$$

Definition 2.24. The proof system SKS_4 is defined to be the six atomic structural rules of SKS, together with $\{\mathsf{s}\downarrow, \mathsf{s}\uparrow, \mathsf{m}, \vee\mathsf{c}\downarrow, \wedge\mathsf{c}\uparrow\}$. The equality rule is restricted to the unit equations $f \vee A = A = t \wedge A$.

Proposition 2.25. *In the presence of the* SKS *equality rule,* $\vee\mathsf{c}\downarrow, \wedge\mathsf{c}\uparrow$ *are derivable for the empty set of rules, and* $\mathsf{s}\downarrow$ *and* $\mathsf{s}\uparrow$ *are all derivable for* $\{\mathsf{s}\}$.

Proof. Derivability of $\vee\mathsf{c}\downarrow$ and $\wedge\mathsf{c}\uparrow$ is clear by associativity and commutativity of $\vee$ and $\wedge$. The rules $\mathsf{s}\downarrow$ and $\mathsf{s}\uparrow$ are each derivable from two applications of the rule s. □

Proposition 2.26. *In a* SKS *derivation, we can replace every instance of* $\vee$ *associativity and commutativity by instances of the rule* $\vee\mathsf{c}\downarrow$ *and unit equality rules (without commutativity we need left and right versions). Dually we can replace every instance of* $\wedge$ *associativity and commutativity by instances of the rule* $\wedge\mathsf{c}\uparrow$ *and unit equality rules.*

In addition, s *is derivable for both* $\mathsf{s}\downarrow$ *and* $\mathsf{s}\uparrow$ *in the presence of either the* SKS *or* SKS_4 *equality rule, i.e. only using unit equations.*

Proof. We perform the following three rewrites, along with their duals, which we omit:

$$
{=}\frac{(A \vee B) \vee C}{A \vee (B \vee C)} \quad\longrightarrow\quad {\vee\mathsf{c}\downarrow}\frac{(A \vee B) \vee {=}\dfrac{C}{\mathsf{f} \vee C}}{{=}\dfrac{A \vee \mathsf{f}}{A} \vee (B \vee C)}
$$

$$
{=}\frac{A \vee B}{B \vee A} \quad\longrightarrow\quad {=}\cfrac{A \vee B}{{\vee\mathsf{c}\downarrow}\cfrac{(\mathsf{f} \vee A) \vee (B \vee \mathsf{f})}{{=}\cfrac{(\mathsf{f} \vee B) \vee (A \vee \mathsf{f})}{B \vee A}}}
$$

$$
{\mathsf{s}}\frac{(A \vee B) \wedge C}{A \vee (B \wedge C)} \quad\longrightarrow\quad {=}\cfrac{(A \vee B) \wedge C}{{\mathsf{s}\downarrow}\cfrac{(A \vee B) \wedge (\mathsf{f} \vee C)}{{=}\cfrac{(A \vee \mathsf{f}) \vee (B \wedge C)}{A \vee (B \wedge C)}}}
$$

□

We could in fact go further, encompassing even atomic rules to this rule shape by employing the *subatomic* methodology [AT17; ATG18]. In what follows, we define a rule, *merge contraction* that has the same behaviour as atomic contractions in decomposition, but corresponds to a generalised notion of a *contractive derivation*. This is in fact true in greater generality for a whole class of nestings, as is proved sub-atomically in [AT17]: those results will be presented and expanded in publications to come, but not in this thesis. Here, our focus is more on the definition and dynamics of merge contractions, especially since they generalise naturally to first-order logic. However, it should be noted that the

$$\mathsf{ai}{\downarrow}\frac{\mathsf{t}}{a \vee \bar{a}} \qquad \mathsf{ai}{\uparrow}\frac{a \wedge \bar{a}}{\mathsf{f}}$$

$$\mathsf{s}{\downarrow}\frac{(A \vee B) \wedge (C \vee D)}{(A \wedge C) \vee (B \vee D)} \qquad \mathsf{s}{\uparrow}\frac{(A \wedge B) \wedge (C \vee D)}{(A \wedge C) \vee (B \wedge D)}$$

$$\mathsf{m}\frac{(A \wedge B) \vee (C \wedge D)}{(A \vee C) \wedge (B \vee D)}$$

$$\mathsf{ac}{\downarrow}\frac{a \vee a}{a} \qquad \mathsf{ac}{\uparrow}\frac{a}{a \wedge a}$$

$$\mathsf{aw}{\downarrow}\frac{\mathsf{f}}{a} \qquad \mathsf{aw}{\uparrow}\frac{a}{\mathsf{t}}$$

$$\vee\mathsf{c}{\downarrow}\frac{(A \vee B) \vee (C \vee D)}{(A \vee C) \vee (B \vee D)} \qquad \wedge\mathsf{c}{\uparrow}\frac{(A \wedge B) \wedge (C \wedge D)}{(A \wedge C) \wedge (B \wedge D)}$$

Figure 2.4: System SKS_4

procedure we are presenting comes directly from the study of decomposition in subatomic systems.

Remark 2.27. Since $\mathsf{s}{\uparrow}$ is derivable for $\{\mathsf{s}\}$ and s for $\{\mathsf{s}{\downarrow}, =\}$, $\mathsf{s}{\uparrow}$ is derivable for $\{\vee\mathsf{c}{\downarrow}, \wedge\mathsf{c}{\uparrow}, \mathsf{s}{\downarrow}\}$. Symmetrically, $\mathsf{s}{\downarrow}$ is derivable for $\{\vee\mathsf{c}{\downarrow}, \wedge\mathsf{c}{\uparrow}, \mathsf{s}{\uparrow}\}$. We choose to include both in the system purely for simplicity and symmetry when presenting the rewriting rules.

In what follows we will present rewriting rules that will allow us to permute critical medials downwards in a derivation. Since instances of medial cannot be permuted with instances of $\mathsf{ac}{\downarrow}$ and $\vee\mathsf{c}{\downarrow}$ that take as premise part of the conclusion of the medial, we introduce the notion of a *(co)contractive derivation*, that bundles together nested instances of medial and instances of $\mathsf{ac}{\downarrow}$ and $\vee\mathsf{c}{\downarrow}$ occurring below it. We also introduce an inference rule, *merge (co)contraction*, that is sound precisely when its premise and conclusion are the premise and conclusion of a (co)contractive derivation. The concept of the *merge* is found in [Gug07], in the context of the BV proof system. Here we adapt and repurpose the notion for SKS_4.

Definition 2.28. The set of *contractive derivations* is the minimal set of $\mathsf{C}_{\downarrow} = \{\mathsf{ac}{\downarrow}, \mathsf{m}, \vee\mathsf{c}{\downarrow}\}$ derivations that satisfy each of the three properties below:

CD1 A formula $A \vee B$ is a contractive derivation.

CD2 $=\frac{\mathsf{f} \vee \mathsf{f}}{\mathsf{f}}, =\frac{\mathsf{t} \vee \mathsf{t}}{\mathsf{t}}$ and $\mathsf{ac}{\downarrow}\frac{a \vee a}{a}$ are contractive derivations.

CD3 $=\frac{\mathsf{f} \vee A}{A}$ and $=\frac{A \vee \mathsf{f}}{A}$ are contractive derivations.

CD4 If $\begin{array}{c} A_1 \vee B_1 \\ \phi_1 \| \mathsf{c}_\downarrow \\ C_1 \end{array}$ and $\begin{array}{c} A_2 \vee B_2 \\ \phi_2 \| \mathsf{c}_\downarrow \\ C_2 \end{array}$ are contractive derivations then both

$$\mathsf{m}\frac{(A_1 \wedge A_2) \vee (B_1 \wedge B_2)}{\boxed{\begin{array}{c} A_1 \vee B_1 \\ \phi_1 \| \mathsf{c}_\downarrow \\ C_1 \end{array}} \wedge \boxed{\begin{array}{c} A_2 \vee B_2 \\ \phi_2 \| \mathsf{c}_\downarrow \\ C_2 \end{array}}} \quad \text{and} \quad \vee\mathsf{c}\!\downarrow\frac{(A_1 \vee A_2) \vee (B_1 \vee B_2)}{\boxed{\begin{array}{c} A_1 \vee B_1 \\ \phi_1 \| \mathsf{c}_\downarrow \\ C_1 \end{array}} \vee \boxed{\begin{array}{c} A_2 \vee B_2 \\ \phi_2 \| \mathsf{c}_\downarrow \\ C_2 \end{array}}}$$

are contractive derivations.

The set of *cocontractive derivations* is defined dually.

Definition 2.29. Given two formulae A, B and $\star \in \{\vee, \wedge\}$, we define their $\star$-*merge set* $M_\star(A,B)$ as the minimum set that satisfies the following conditions

M1 For any A and B, $A \star B \in M_\star(A,B)$.

M2 For any atom or unit a, $a \in M_\star(a,a)$.

M3 For any A, $A \in M_\vee(A,\mathsf{f}), A \in M_\vee(\mathsf{f},A), A \in M_\wedge(A,\mathsf{t})$ and $A \in M_\wedge(\mathsf{t},A)$.

M4 For $\alpha \in \{\vee, \wedge\}$, if $C_1 \in M_\star(A_1,B_1)$ and $C_2 \in M_\star(A_2,B_2)$, then

$$C_1 \ \alpha \ C_2 \in M_\star(A_1 \ \alpha \ A_2, B_1 \ \alpha \ B_2)$$

.

If $C \in M_\star(A,B)$, we say C is a $\star$-*merge* of A and B. We call

$$M^P_\star(A,B) ;= M_\star(A,B) \backslash \{A \star B\}$$

the *proper* $\star$*-merge set* of A and B. If $C \in M^P_\star(A,B)$, then we say C is a *non-trivial* $\star$*-merge* of A and B.

Proposition 2.30. *There exists a contractive derivation* $\begin{array}{c} A \vee B \\ \phi \| \mathsf{c}_\downarrow \\ C \end{array}$ *iff* $C \in M_\vee(A,B)$.

Dually there exists a cocontractive derivation $\begin{array}{c} C \\ \phi \| \mathsf{c}_\uparrow \\ A \wedge B \end{array}$ *iff* $C \in M_\wedge(A,B)$..

Proof. We first prove the left to right direction by induction on the size of ϕ.

- Clearly **M1** covers every contractive derivation generated by **CD1**, **M2** covers every contractive derivation generated by **CD2** and **M3** covers every contractive derivation generated by **CD3**.
- Let ϕ be a contractive derivation with at least one inference rule that is not ac$\downarrow$ or $=$. By **CD4** it must be of the form

$$\mathsf{m}\frac{(A_1 \wedge A_2) \vee (B_1 \wedge B_2)}{\boxed{\begin{array}{c} A_1 \vee B_1 \\ \phi_1 \| \mathsf{c}_\downarrow \\ C_1 \end{array}} \wedge \boxed{\begin{array}{c} A_2 \vee B_2 \\ \phi_2 \| \mathsf{c}_\downarrow \\ C_2 \end{array}}} \quad \text{or} \quad \vee\mathsf{c}\!\downarrow\frac{(A_1 \vee A_2) \vee (B_1 \vee B_2)}{\boxed{\begin{array}{c} A_1 \vee B_1 \\ \phi_1 \| \mathsf{c}_\downarrow \\ C_1 \end{array}} \vee \boxed{\begin{array}{c} A_2 \vee B_2 \\ \phi_2 \| \mathsf{c}_\downarrow \\ C_2 \end{array}}}$$

with ϕ_1 and ϕ_2 contractive derivations. By the inductive hypothesis, $C_1 \in M_\vee(A_1, B_1)$ and $C_2 \in M_\vee(A_2, B_2)$. Therefore, in both cases $C = (C_1\ \alpha\ C_2) \in M_\vee(A,B) = M_\vee(A_1\ \alpha\ A_2, B_1\ \alpha\ B_2)$ by **M4**.

To prove the right to left direction we proceed by structural induction on C:

- If C is an atom or unit a, then a must be in $M_\vee(A,B)$ by virtue of **M2**, so it must be the case that $A \equiv B \equiv a$. Therefore, by **CD2**, $\phi \equiv \mathsf{c}{\downarrow}\dfrac{a \vee a}{a}$ is a suitable contractive derivation.
- If C is not an atom, then either $C \equiv [A \vee B]$ and is in the $\vee$-merge set by virtue of **M1**, in which case we take $\phi \equiv A \vee B$ as our contractive derivation, by **CD1**. C could also be in the merge set by virtue of **M3**, in which case we can take $\phi \equiv {=}\dfrac{\mathsf{f} \vee A}{A}$ or $\phi \equiv {=}\dfrac{A \vee \mathsf{f}}{A}$, by **CD3**. If it is neither of these, then it must be the case that $C \equiv [C_1\ \alpha\ C_2]$, with $A \equiv A_1\ \alpha\ A_2$, $B \equiv B_1\ \alpha\ B_2$ where $\alpha \in \{\vee, \wedge\}$ and $C_1 \in M_\vee(A_1, B_1)$ and $C_2 \in M_\vee(A_2, B_2)$. By the inductive hypothesis, there are contractive derivations $\begin{array}{c} A_1 \vee B_1 \\ \phi_1 \| \mathsf{c}_\downarrow \\ C_1 \end{array}$ and $\begin{array}{c} A_2 \vee B_2 \\ \phi_2 \| \mathsf{c}_\downarrow \\ C_2 \end{array}$. We can therefore combine these with m or $\vee\mathsf{c}{\downarrow}$ to form:

$$\phi \equiv \mathsf{m}\dfrac{(A_1 \wedge A_2) \vee (B_1 \wedge B_2)}{\boxed{\begin{array}{c} A_1 \vee B_1 \\ \phi_1 \| \mathsf{c}_\downarrow \\ C_1 \end{array}} \wedge \boxed{\begin{array}{c} A_2 \vee B_2 \\ \phi_2 \| \mathsf{c}_\downarrow \\ C_2 \end{array}}} \quad \text{or} \quad \phi \equiv \vee\mathsf{c}{\downarrow}\dfrac{(A_1 \vee A_2) \vee (B_1 \vee B_2)}{\boxed{\begin{array}{c} A_1 \vee B_1 \\ \phi_1 \| \mathsf{c}_\downarrow \\ C_1 \end{array}} \vee \boxed{\begin{array}{c} A_2 \vee B_2 \\ \phi_2 \| \mathsf{c}_\downarrow \\ C_2 \end{array}}}$$

  which are contractive derivations by **CD4**.

We prove the dual likewise. □

Corollary 2.31. *$C \in M_\vee(A,B)$ iff $\bar{C} \in M_\wedge(\bar{A}, \bar{B})$*

Proof. Immediate from the fact that each contractive derivation is clearly the negation of a cocontractive derivation and vice versa. □

Given the above characterisation of contractive derivations where the conclusion is a $\vee$-merge of the premise, we can characterise contractive derivations with a single inference rule, merge (co)contraction.

Definition 2.32. $\mathsf{mc}{\downarrow}\dfrac{A \vee B}{C}$ is a *merge contraction* if $C \in M^P_\vee(A,B)$. $\mathsf{mc}{\uparrow}\dfrac{C}{A \wedge B}$ is a *merge cocontraction* if $C \in M^P_\wedge(A,B)$. If $\mathsf{mc}{\downarrow}\dfrac{A \vee B}{C}$ is an instance of merge contraction, we call the minimal contractive derivation $\begin{array}{c} A \vee B \\ \phi \| \mathsf{c}_\downarrow \\ C \end{array}$ guaranteed by Proposition 2.30 the *associated derivation* to $\mathsf{mc}{\downarrow}\dfrac{A \vee B}{C}$.

Remark 2.33. As $A \in M^P_\vee(A,A)$, general contractions $\mathsf{c}\!\downarrow \frac{A \vee A}{A}$ are a special instance of merge contractions, i.e., for any formula A there is a merge contraction

$$\mathsf{mc}\!\downarrow \frac{A \vee A}{A}$$

and its associated derivation

$$\begin{array}{c} A \vee A \\ \phi \,\| \, \mathsf{c}_\downarrow \\ A \end{array}$$

The size of the associated derivations is quadratic in the size of A.

Every instance of medial is also an instance of both a merge contraction and a merge cocontraction by **M1** and **M4**, with the associated derivations being the medials themselves. Similarly, every instance $\vee\mathsf{c}\!\downarrow$ is also an instance of a merge contraction, and every instance of $\wedge\mathsf{c}\!\uparrow$ is also an instance of a merge cocontraction, also by **M1** and **M4**.

Definition 2.34. We define SKSm to be:

$$(\mathsf{SKS}_4 \backslash \{\mathsf{ac}\!\downarrow, \mathsf{ac}\!\uparrow, \mathsf{m}, \vee\mathsf{c}\!\downarrow, \wedge\mathsf{c}\!\uparrow\}) \cup \{\mathsf{mc}\!\downarrow, \mathsf{mc}\!\uparrow\}$$

as shown in Figure 2.5. There is a map $\mu : \mathsf{SKSm} \to \mathsf{SKS}_4$, defined such that every merge contraction maps onto its associated derivation. There are two obvious maps from SKS_4 to SKSm: $\nu_{\min} : \mathsf{SKS}_4 \to \mathsf{SKSm}$, which maps every instance of $\mathsf{ac}\!\downarrow$, m and $\vee\mathsf{c}\!\downarrow$ each to a single instance of $\mathsf{mc}\!\downarrow$ and $\nu_{\max} : \mathsf{SKS}_4 \to$ SKSm which maps maximal contractive derivations to merge contractions.

As can be seen in the example below, μ is the left inverse map to both $\nu_{\min}$ and $\nu_{\max}$. It would seem as if one could construe $\nu_{\min}$ and $\nu_{\max}$ as the left and right adjoint functors to μ, but this is beyond the scope of this thesis.

Convention 2.35. From now on, since the translation between the two proof systems is so straightforward, we will often just refer to SKS when, strictly, we should refer to SKS_4.

There is one final way we can characterise a merge contraction: as a contraction on contexts.

Definition 2.36. If $K\{\ \}$ is a context with n-holes, and $A_1, \ldots, A_n, B_1, \ldots, B_n$ are formulas, then

$$\mathsf{cc}\!\downarrow \frac{K\{A_1\}\ldots\{A_n\} \vee K\{B_1\}\ldots\{B_n\}}{K\{A_1 \vee B_1\}\ldots\{A_n \vee B_n\}} \quad \text{and} \quad \mathsf{cc}\!\uparrow \frac{K\{A_1 \wedge B_1\}\ldots\{A_n \wedge B_n\}}{K\{A_1\}\ldots\{A_n\} \wedge K\{B_1\}\ldots\{B_n\}}$$

are instances of inference rules, called *context contraction* and *context cocontraction* respectively.

First, we show that every context contraction is also a merge contraction.

Proposition 2.37. *If*

$$\mathsf{cc}\!\downarrow \frac{K\{A_1\}\ldots\{A_n\} \vee K\{B_1\}\ldots\{B_n\}}{K\{A_1 \vee B_1\}\ldots\{A_n \vee B_n\}}$$

$\mathsf{ai}\!\downarrow \dfrac{\mathsf{t}}{a \vee \bar{a}}$ identity	$\mathsf{ai}\!\uparrow \dfrac{a \wedge \bar{a}}{\mathsf{f}}$ cut
$\mathsf{aw}\!\downarrow \dfrac{\mathsf{f}}{a}$ weakening	$\mathsf{aw}\!\uparrow \dfrac{a}{\mathsf{t}}$ coweakening
$\mathsf{mc}\!\downarrow \dfrac{A \vee B}{C}$ merge contraction	$\mathsf{mc}\!\uparrow \dfrac{C}{A \wedge B}$ merge cocontraction
$\mathsf{s}\!\downarrow \dfrac{(A \vee B) \wedge (C \vee D)}{(A \wedge C) \vee (B \vee D)}$ switch (down)	$\mathsf{s}\!\uparrow \dfrac{(A \wedge B) \wedge (C \vee D)}{(A \wedge C) \vee (B \wedge D)}$ switch (up)

Figure 2.5: System SKSm

is a valid instance of a context contraction then

$$\mathsf{mc}\!\downarrow \frac{K\{A_1\}\ldots\{A_n\} \vee K\{B_1\}\ldots\{B_n\}}{K\{A_1 \vee B_1\}\ldots\{A_n \vee B_n\}}$$

is a valid instance of a merge contraction.

Proof. By structural induction on $K\{\ \}$. If $K\{\ \} \equiv \{\ \}$, then by **M1**, $\mathsf{mc}\!\downarrow \dfrac{A_1 \vee B_1}{A_1 \vee B_1}$ is a valid merge contraction. If $K\{\ \} \equiv a$, for some atom a, then by **M2**, $\mathsf{mc}\!\downarrow \dfrac{a \vee a}{a}$ is valid merge contraction. Now assume

$$\mathsf{cc}\!\downarrow \frac{K_1\{A_1\}\ldots\{A_k\} \vee K_1\{B_1\}\ldots\{B_k\}}{K_1\{A_1 \vee B_1\}\ldots\{A_k \vee B_k\}} \quad \text{and} \quad \mathsf{cc}\!\downarrow \frac{K_2\{A_{k+1}\}\ldots\{A_n\} \vee K_2\{B_{k+1}\}\ldots\{B_n\}}{K_2\{A_{k+1} \vee B_{k+1}\}\ldots\{A_n \vee B_n\}}$$

are valid instances of context contractions. By the IH, they are also instances of merge contractions. Thus, by **M4**,

$$\mathsf{mc}\!\downarrow \frac{(K_1\{A_1\}\ldots\{A_k\}\ \alpha\ K_2\{A_{k+1}\}\ldots\{A_n\}) \vee (K_1\{B_1\}\ldots\{B_k\}\ \alpha\ K_2\{B_{k+1}\}\ldots\{B_n\})}{K_1\{(A_1 \vee B_1)\}\ldots\{(A_k \vee B_k)\}\ \alpha\ K_2\{(A_{k+1} \vee B_{k+1})\}\ldots\{(A_n \vee B_n)\}}$$

is a valid instance of a merge contraction, for $\alpha \in \{\wedge, \vee\}$. □

Now, we show that every merge contraction is a context contration. However the natural definition for context contractions, which we have given above, does not quite match up to our merge contraction definition: they are equivalent to merge contractions with out the **M3** rule. However, we only need to add a minor condition to get a working equivalence between the two inference rules.

Proposition 2.38. *If* $\mathsf{mc}\!\downarrow \dfrac{A \vee B}{C}$ *is a valid instance of a merge contraction, then there is some n s.t. we can find a context $K\{\ \}\ldots\{\ \}$ and formulae $A_1, \ldots, A_n, B_1, \ldots, B_n$ s.t. $A \equiv K\{A_1\}\ldots\{A_n\}$, $B \equiv K\{B_1\}\ldots\{B_n\}$,*

$$\mathsf{cc}\!\downarrow \frac{K\{A_1\}\ldots\{A_n\} \vee K\{B_1\}\ldots\{B_n\}}{K\{A_1 \vee B_1\}\ldots\{A_n \vee B_n\}}$$

is a valid instance of a context contraction, and $C \equiv K\{C_1\}\ldots\{C_n\}$, where either $C_i \equiv A_i \vee B_i$, $C_i \equiv A_i$ and $B_i \equiv \mathsf{f}$, or $C_i \equiv B_i$ and $A_i \equiv \mathsf{f}$.

Proof. We proceed by case analysis on $\mathsf{mc}\!\downarrow \dfrac{A \vee B}{C}$. If $C \in M_\vee(A,B)$ due to **M1**, then we choose $K\{\ \} = \{\ \}$. If $C \in M_\vee(A,B)$ due to **M2**, then we choose $K\{\ \} = a$. If $C \in M_\vee(A,B)$ due to **M3**, then the condition on units means that we can choose $K\{\ \} = \{\ \}$. If $C \in M_\vee(A,B)$ due to **M4**, then we proceed by induction on the size of A. The base cases consist of the previous three cases. Assume now that $A = [A_1\ \alpha\ A_2]$, $B = [B_1\ \alpha\ B_2]$ and $C = [C_1\ \alpha\ C_2]$, with $C_1 \in M_\vee(A_1, B_1)$ and $C_2 \in M_\vee(A_2, B_2)$. Then, by the IH, we have $A_1 \equiv K_1\{A_{11}\}\ldots\{A_{1m}\}$, $B_1 \equiv K_1\{B_{11}\}\ldots\{B_{1m}\}$, $C_1 \equiv K_1\{C_{11}\}\ldots\{C_{1m}\}$, $A_1 \equiv K_2\{A_{21}\}\ldots\{A_{2n}\}$, $B_2 \equiv K_2\{B_{21}\}\ldots\{B_{2n}\}$, $C_2 \equiv K_2\{C_{21}\}\ldots\{C_{2n}\}$ where the C_i are as in proposition statement. Thus, we can take $K\{\ \}\ldots\{\ \} \equiv K_1\{\ \}\ldots\{\ \}\ \alpha\ K_2\{\ \}\ldots\{\ \}$, which is a context with $m+n$ holes. Then

$$\mathsf{cc}\!\downarrow \frac{K\{A_{11}\}\ldots\{A_{1m}\}\{A_{21}\}\ldots\{A_{2n}\} \vee K\{B_{11}\}\ldots\{B_{1m}\}\{B_{21}\}\ldots\{B_{2n}\}}{K\{A_{11} \vee B_{11}\}\ldots\{A_{1m} \vee B_{1m}\}\{A_{21} \vee B_{21}\}\ldots\{A_{2n} \vee B_{2n}\}}$$

is a valid instance of a context contraction. □

Example 2.39. In the center below is a SKS derivation, with a SKSm derivation either side. The maps $\mu, \nu_{\min}$ and $\nu_{\max}$ between them are shown. Note that the derivation in the center is the associated derivation to the merge contraction on the left.

$$\begin{tikzcd}
\mathsf{mc}\!\downarrow \dfrac{(a \wedge b) \vee (a \wedge c)}{a \wedge (b \vee c)} \arrow[r, bend right, "\mu"'] &
\mathsf{m} \dfrac{(a \wedge b) \vee (a \wedge c)}{\mathsf{ac}\!\downarrow \dfrac{a \vee a}{a} \wedge (b \vee c)} \arrow[l, bend right, "\nu_{\max}"'] \arrow[r, bend left, "\nu_{\min}"] &
\mathsf{mc}\!\downarrow \dfrac{(a \wedge b) \vee (a \wedge c)}{\mathsf{mc}\!\downarrow \dfrac{a \vee a}{a} \wedge (b \vee c)} \arrow[l, bend left, "\mu"]
\end{tikzcd}$$

We can rewrite the derivations on the left and right with context contractions instead of merge contractions, where $K_1\{\ \} = a \wedge \{\ \}$, $K_2\{\ \}\{\ \} = \{\ \} \wedge \{\ \}$ and $K_3 = a$

$$\mathsf{cc}\!\downarrow \frac{K_1\{b\} \vee K_1\{c\}}{K_1\{b \vee c\}} \quad \text{and} \quad \mathsf{cc}\!\downarrow \frac{K_2\{a\}\{b\} \vee K_2\{a\}\{c\}}{K_2\left\{\mathsf{cc}\!\downarrow \dfrac{K_3 \vee K_3}{K_3}\right\}\{b \vee c\}}$$

Any proposition in the merge set $M_\vee(A,B)$ can be derived from both A and B, and in fact we can define a canonical derivation of just weakenings that does so, which we call the projective derivations associated to $M_\vee(A,B)$.

Definition 2.40. If $C \in M_\vee(A,B)$, we define the *projective derivations* $\begin{array}{c} A \\ \pi_A \Vert \{\mathsf{aw}\downarrow\} \\ C \end{array}$ *and* $\begin{array}{c} B \\ \pi_B \Vert \{\mathsf{aw}\downarrow\} \\ C \end{array}$ *associated to* $M_\vee(A,B)$ as follows:

- If $C \equiv A \vee B$, we take $\pi_A \equiv {=}\dfrac{A}{\begin{array}{c} \\ A \vee \end{array}\begin{array}{c} \mathsf{f} \\ \Vert \{\mathsf{aw}\downarrow\} \\ B \end{array}}$ and $\pi_B \equiv {=}\dfrac{B}{\begin{array}{c} \mathsf{f} \\ \Vert \{\mathsf{aw}\downarrow\} \\ A \end{array}\begin{array}{c} \\ \vee B \end{array}}$
- If $C \equiv A \equiv B \equiv a$ with a a unit or an atom, we take $\pi_A \equiv a$ and $\pi_B \equiv a$.
- If $A \equiv C$ and $B \equiv \mathsf{f}$, then we take $\pi_A \equiv A$ and $\pi_B \equiv \begin{array}{c} \mathsf{f} \\ \Vert \{\mathsf{aw}\downarrow\} \\ B \end{array}$, and vice versa if $A \equiv \mathsf{f}$ and $B \equiv C$.
- If $C \equiv C_1 \; \alpha \; C_2$, $A \equiv A_1 \; \alpha \; A_2$, $B \equiv B_1 \; \alpha \; B_2$ where $\alpha \in \{\vee, \wedge\}$ and $C_1 \in M_\vee(A_1,B_1)$ and $C_2 \in M_\vee(A_2,B_2)$, we take

$$\psi_A \equiv \boxed{\begin{array}{c} A_1 \\ \pi_{A_1} \Vert \{\mathsf{aw}\downarrow\} \\ C_1 \end{array}} \; \alpha \; \boxed{\begin{array}{c} A_2 \\ \pi_{A_2} \Vert \{\mathsf{aw}\downarrow\} \\ C_2 \end{array}} \quad \text{and} \quad \psi_B \equiv \boxed{\begin{array}{c} B_1 \\ \pi_{B_1} \Vert \{\mathsf{aw}\downarrow\} \\ C_1 \end{array}} \; \alpha \; \boxed{\begin{array}{c} B_2 \\ \pi_{B_2} \Vert \{\mathsf{aw}\downarrow\} \\ C_2 \end{array}} \quad ,$$

  where π_{A_1}, π_{B_1} are the projective derivations associated to $M_\vee(A_1,B_1)$ and π_{A_2}, π_{B_2} are the projective definitions associated to $M_\vee(A_2,B_2)$.

2.3 Rewriting Systems for Merge Contractions

We will now present reduction rules to permute merge contraction instances downwards in a SKSm proof. By having associativity and commutativity fit into a single rule scheme and uniting all contractive rules (m, $\vee\mathsf{c}\downarrow$ and $\mathsf{ac}\downarrow$) as merge contractions, we greatly reduce the ways in which rules can overlap, leaving us with very few rewrites to define. Of course, we could translate these rewrites into SKS or SKS_4, but doing so would be at the cost of many more cases to consider in the rewrites and a much more complex termination proof.

The general nature of these reduction rules means that they can be defined for a variety of systems to permute a variety of rules [AT17], and fit into a wider normalisation theory [ATG18]. A general study on decomposition utilising merge contractions will be provided in publications to come.

2.3.1 Three useful lemmas involving merges

First, we define a precise notion of subformula for merges, to avoid counting "accidental" subformulae later on. We will then state and prove three useful lemmas involving merges, which will allow us to permute merge contractions past other rules more simply.

Definition 2.41. Let $K\{D\} \in M_\vee(A,B)$. We say that D is a *K-subformula* in A or B in the following cases:

M1 If $K\{D\} \equiv K'\{D\} \vee B$ and $A \equiv K'\{D\}$ then D is a K-subformula of A. If $K\{D\} \equiv A \vee K'\{D\}$ and $B \equiv K'\{D\}$ then D is a K-subformula of B.

M3 If $K\{D\} \equiv A$, then D is a K-subformula of A.

M4 If $K_1\{D_1\} \in M_\vee(A_1,B_1)$ and D_1 is a K_1-subformula of A_1 (resp. B_1), then for all $C_2 \in M_\vee(A_2,B_2)$, D_1 is a K-subformula of $A_1 \;\alpha\; A_2$ (resp. $B_1 \;\alpha\; B_2$), where $K\{\ \} = K_1\{\ \} \;\alpha\; C_2$.

Remark 2.42. It should be straightforward to observe that if D is a K-subformula of A, then it is a subformula of A. The notion of K-subformula is essentially that of a subformula occurence, but we redefine it here for those not familiar with the notion.

Lemma 2.43. *Let A,B,C be formulae with $C \in M^P_\vee(A,B)$. Let $K_C\{\ \}$ and P_C be s.t. $C \equiv K_C\{P_C\}$ and P_C is not a K_C-subformula of A or B. Then we can find contexts $K_A\{\ \}, K_B\{\ \}$ and formulae P_A, P_B s.t. A factorises as $K_A\{P_A\}$, B as $K_B\{P_B\}$, $P_C \in M_\vee(P_A,P_B)$ and for any Q_A, Q_B, if $Q_C \in M_\vee(Q_A,Q_B)$, then*

$$K_C\{Q_C\} \in M_\vee(K_A\{Q_A\}, K_B\{Q_B\}) \quad .$$

Alternatively, we can always find $K_A\{\ \}$, $K_B\{\ \}$, P_A and P_B s.t.

$$\mathsf{mc}{\downarrow}\frac{A \vee B}{K_C\{P_C\}} \quad \equiv \quad \mathsf{mc}{\downarrow}\frac{K_A\{P_A\} \vee K_B\{P_B\}}{K_C\{P_C\}}$$

and, for any $Q_C \in M_\vee(Q_A,Q_B)$,

$$\mathsf{mc}{\downarrow}\frac{K_A\{Q_A\} \vee K_B\{Q_B\}}{K_C\{Q_C\}}$$

is a valid instance of $\mathsf{mc}{\downarrow}$.

Proof. We proceed by induction on the size of $K_C\{\ \}$. If $K_C\{\ \} = \{\ \}$, then we have $K_A\{\ \} = K_B\{\ \} = \{\ \}, P_A \equiv A$ and $P_B \equiv B$.

Now assume $K_C\{\ \} = D \;\alpha\; K'_C\{\ \}$. Since $A \neq \mathsf{f} \neq B$ (as P_C is not a K_C-subformula of A or B) and $C \neq A \vee B$, we must have that $C \equiv D \;\alpha\; K'_C\{P_C\} \in M^P_\vee(A,B)$ by virtue of **M4**, with $A \equiv A_1 \;\alpha\; A_2$, $B \equiv B_1 \;\alpha\; B_2$, $D \in M_\vee(A_1,B_1)$ and $K'_C\{P_C\} \in M_\vee(A_2,B_2)$.

If $K'_C\{P_C\} \equiv (A_2 \vee B_2)$ and $P_C \neq (A_2 \vee B_2)$ then clearly P_C is a K'_C-subformula of A_2 (or B_2) and therefore also a K_C-subformula of A (or B). Therefore, if $K'_C\{P_C\} \equiv (A_2 \vee B_2)$ then $P_C \equiv (A_2 \vee B_2)$, and we can factorise A as $A_1 \;\alpha\; \{A_2\}$ and B as $B_1 \;\alpha\; \{B_2\}$. It is straightforward to check that all the conditions hold.

If we have that $K'_C\{P_C\} \neq (A_2 \vee B_2)$, then $K'_C\{P_C\} \in M^P_\vee(A_2,B_2)$ and, by the IH, we have $A_2 \equiv K_{A_2}\{P_A\}$ and $B_2 \equiv K_{B_2}\{P_B\}$ with all the appropriate conditions. Thus, we can factorise A as $A_1 \;\alpha\; K_{A_2}\{P_A\}$ and B as $B_1 \;\alpha\; K_{B_2}\{P_B\}$. By the IH, $P_C \in M_\vee(P_A,P_B)$.

Let $Q_C \in M_\vee(Q_A,Q_B)$. By the IH, $K'_C\{Q_C\} \in M_\vee(K_{A_2}\{Q_A\}, K_{B_2}\{Q_B\})$. So $K_C\{Q_C\} \equiv D \;\alpha\; K'_C\{Q_C\} \in M_\vee(A_1 \;\alpha\; K_{A_2}\{Q_A\}, B_1 \;\alpha\; K_{B_2}\{Q_B\}) \equiv M_\vee(K_A\{Q_A\}, K_B\{Q_B\})$. □

Lemma 2.44. *Assume we have $A \in M_\vee(A_1, A_2), B \in M_\vee(B_1, B_2)$ and $C \in M^P_\vee(A, B)$. Then we can find C_1, C_2 s.t. $C_1 \in M^P_\vee(A_1, B_1)$ and $C_2 \in M^P_\vee(A_2, B_2)$ s.t. $C \in M^P_\vee(C_1, C_2)$. Alternatively, we always can find C_1 and C_2 s.t. the following rewrite is valid.*

$$\mathsf{mc}{\downarrow}\frac{\mathsf{mc}{\downarrow}\dfrac{(A_1 \vee B_1) \vee (A_2 \vee B_2)}{A \vee B}}{C} \quad\longrightarrow\quad \mathsf{mc}{\downarrow}\frac{\mathsf{mc}{\downarrow}\dfrac{A_1 \vee B_1}{C_1} \vee \mathsf{mc}{\downarrow}\dfrac{A_2 \vee B_2}{C_2}}{C}$$

Proof. We proceed by structural induction on C.

If $C \equiv a$, where a is an atom or unit, then clearly $A_1, A_2, B_1, B_2 \in \{a, \mathsf{f}\}$, with at least one of the four not equal to f if $a \neq \mathsf{f}$. Thus it is straightforward to appropriately choose C_1, C_2 from $\{a, \mathsf{f}\}$.

If $C \equiv D_1 \; \alpha \; D_2$, then either:

M3 $A \equiv D_1 \; \alpha \; D_2$, $B \equiv \mathsf{f}$ (WLOG). Clearly, $B_1 \equiv B_2 \equiv \mathsf{f}$. Thus we can take $C_1 \equiv A_1$ and $C_2 \equiv A_2$, since $D_1 \; \alpha \; D_2 \in M^P_\vee(A_1, A_2)$.

M4 $A \equiv E_1 \; \alpha \; E_2$, $B \equiv F_1 \; \alpha \; F_2$ with $D_1 \in M_\vee(E_1, F_1)$ and $D_2 \in M_\vee(E_2, F_2)$. There are lots of different permutations possible at this point, but they boil down to 4 fundamentally different cases:

- $A_1 \equiv E_1 \; \alpha \; E_2$, $B_1 \equiv F_1 \; \alpha \; F_2$ and $A_2 \equiv B_2 \equiv \mathsf{f}$. In this case, we can set $C_1 \equiv D_1 \; \alpha \; D_2$ and $C_2 \equiv \mathsf{f}$. Symmetrically, if $A_1 \equiv B_1 \equiv \mathsf{f}$, $A_2 \equiv E_1 \; \alpha \; E_2$ and $B_2 \equiv F_1 \; \alpha \; F_2$, then we can set $C_1 \equiv \mathsf{f}$ and $C_2 \equiv D_1 \; \alpha \; D_2$.
- $A_1 \equiv E_1 \; \alpha \; E_2$, $B_1 \equiv \mathsf{f}$, $A_2 \equiv \mathsf{f}$ and $B_2 \equiv F_1 \; \alpha \; F_2$ or $A_1 \equiv \mathsf{f}$, $B_1 \equiv F_1 \; \alpha \; F_2$, $A_2 \equiv E_1 \; \alpha \; E_2$ and $B_2 \equiv \mathsf{f}$. In both these cases, we can set $C_1 \equiv E_1 \; \alpha \; E_2$, $C_2 \equiv F_1 \; \alpha \; F_2$.
- $A_1 \equiv A_{11} \; \alpha \; A_{12}$, $B_1 \equiv F_1 \; \alpha \; F_2$, $A_2 \equiv A_{21} \; \alpha \; A_{22}$ and $B_2 \equiv \mathsf{f} \; \alpha \; \mathsf{f}$ (B_2 is chosen as $\mathsf{f} \; \alpha \; \mathsf{f}$ not f for a technical reason, it is not important). It must be the case that $C_1 \equiv C_{11} \; \alpha \; C_{12}$ and $C_2 \equiv C_{21} \; \alpha \; C_{22}$ for some $C_{11}, C_{12}, C_{21}, C_{22}$, s.t. $C_{11} \in M_\vee(A_{11}, F_1)$, $C_{12} \in M_\vee(A_{12}, F_2)$, $C_{21} \in M_\vee(A_{21}, \mathsf{f})$, $C_{22} \in M_\vee(A_{22}, \mathsf{f})$, $D_1 \in M_\vee(C_{11}, C_{21})$, and $D_2 \in M_\vee(C_{12}, C_{22})$. We will show how to find C_{11} and C_{21}; finding C_{12} and C_{22} is analogous.
  - If $D_1 \equiv E_1 \vee F_1$, then we can take $C_{11} \equiv A_{11} \vee F_1$ and $C_{12} \equiv A_{21} \vee \mathsf{f}$.
  - If $D_1 \neq E_1 \vee F_1$, then $D_1 \in M^P_\vee(E_1, F_1)$ and, by the IH, we can find $G_1 \in M^P_\vee(A_{11}, F_1)$ and $G_2 \in M^P_\vee(A_{21}, \mathsf{f})$ with $D_1 \in M^P_\vee(G_1, G_2)$. Clearly we can set $C_{11} \equiv G_1$ and $C_{21} \equiv G_2$.

  Any other case where exactly one of A_1, B_1, A_2, B_2 is equal to f (or $\mathsf{f} \; \alpha \; \mathsf{f}$) is analogous.
- $A_1 \equiv A_{11} \; \alpha \; A_{12}$, $B_1 \equiv B_{11} \; \alpha \; B_{12}$, $A_2 \equiv A_{21} \; \alpha \; A_{22}$ and $B_1 \equiv B_{21} \; \alpha \; B_{22}$. It must be the case the $C_1 \equiv C_{11} \; \alpha \; C_{12}$ and $C_2 \equiv C_{21} \; \alpha \; C_{22}$ for some $C_{11}, C_{12}, C_{21}, C_{22}$, s.t. $C_{11} \in M_\vee(A_{11}, B_{11})$, $C_{12} \in M_\vee(A_{12}, B_{12})$, $C_{21} \in M_\vee(A_{21}, B_{21})$, $C_{22} \in M_\vee(A_{22}, B_{22})$, $D_1 \in M_\vee(C_{11}, C_{21})$ and $D_2 \in$

$M_\vee(C_{12}, C_{22})$. We will show how to find C_{11} and C_{21}, finding C_{12} and C_{22} is analogous.

- If $D_1 \equiv E_1 \vee F_1$, then we can take $C_{11} \equiv A_{11} \vee B_{11}$ and $C_{12} \equiv A_{21} \vee B_{21}$.
- If $D_1 \not\equiv E_1 \vee F_1$, then $D_1 \in M^P_\vee(E_1, F_1)$ and, by the IH, we can find $G_1 \in M^P_\vee(A_{11}, B_{11})$ and $G_2 \in M^P_\vee(A_{21}, B_{21})$ with $D_1 \in M^P_\vee(G_1, G_2)$. Clearly we can set $C_{11} \equiv G_1$ and $C_{21} \equiv G_2$.

□

Lemma 2.45. *Assume we have $B \in M^P_\vee(A_1, A_2)$ and $B \in M^P_\wedge(C_1, C_2)$. Then we can find B_1, B_2, B_3, B_4 s.t. $A_1 \in M^P_\wedge(B_1, B_2)$, $A_2 \in M^P_\wedge(B_3, B_4)$, $C_1 \in M^P_\vee(B_1, B_3)$ and $C_2 \in M^P_\vee(B_2, B_4)$. Alternatively, we can always find B_1, B_2, B_3, B_4 s.t. the following rewrite is valid (although in practice the medial would be absorbed into one of the merge (co)contractions):*

$$
\mathsf{mc}\!\uparrow \frac{\mathsf{mc}\!\downarrow \dfrac{A_1 \vee A_2}{B}}{C_1 \wedge C_2}
\quad\longrightarrow\quad
\mathsf{m}\,\frac{\mathsf{mc}\!\uparrow \dfrac{A_1}{B_1 \wedge B_2} \vee \mathsf{mc}\!\uparrow \dfrac{A_2}{B_3 \wedge B_4}}{\mathsf{mc}\!\downarrow \dfrac{B_1 \vee B_3}{C_1} \wedge \mathsf{mc}\!\downarrow \dfrac{B_2 \vee B_4}{C_2}}
$$

Proof. We proceed by structural induction on B.

If $B \equiv a$, where a is an atom or unit, then clearly $A_1, A_2, C_1, C_2 \in \{a, \mathsf{f}, \mathsf{t}\}$. It is straightforward to select appropriate B_i from $\{a, \mathsf{f}, \mathsf{t}\}$.

If $B \equiv D_1 \; \alpha \; D_2$, then either:

M3$_A$ $A_1 \equiv D_1 \; \alpha \; D_2$ and $A_2 \equiv \mathsf{f}$ (WLOG). Then we can set $B_1 \equiv C_1$, $B_2 \equiv C_2$, $B_3 \equiv \mathsf{f}$ and $B_4 \equiv \mathsf{f}$.

M3$_C$ $C_1 \equiv D_1 \; \alpha \; D_2$ and $C_2 \equiv \mathsf{t}$ (WLOG). Then we can set $B_1 \equiv A_1$, $B_2 \equiv \mathsf{t}$, $B_3 \equiv A_2$ and $B_4 \equiv \mathsf{t}$.

M4 $A_1 \equiv (A_{11} \; \alpha \; A_{12})$, $A_2 \equiv (A_{21} \; \alpha \; A_{22})$, $C_1 \equiv (C_{11} \; \alpha \; C_{12})$ and $C_2 \equiv (C_{21} \; \alpha \; C_{22})$, with $D_1 \in M_\vee(A_{11}, A_{21})$, $D_2 \in M_\vee(A_{12}, A_{22})$, $D_1 \in M_\wedge(C_{11}, C_{21})$ and $D_2 \in M_\wedge(C_{12}, C_{22})$.

Either $D_1 \equiv A_{11} \vee A_{21}, D_1 \equiv C_{11} \wedge C_{21}$ or $D_1 \in M^P_\vee(A_{11}, A_{21}) \cap M^P_\wedge(C_{11}, C_{21})$.

- If $D_1 \equiv A_{11} \vee A_{21}$, then $C_{11} \equiv C_{111} \vee C_{112}$ and $C_{21} \equiv C_{211} \vee C_{212}$, with $A_{11} \in M_\wedge(C_{111}, C_{211})$ and $A_{21} \in M_\wedge(C_{112}, C_{212})$.
- If $D_1 \equiv C_{11} \wedge C_{21}$, then $A_{11} \equiv A_{111} \wedge A_{112}$ and $A_{21} \equiv A_{211} \wedge A_{212}$ with $C_{11} \in M_\vee(A_{111}, A_{211})$ and $C_{21} \in M_\vee(A_{112}, A_{212})$.
- If $D_1 \in M^P_\vee(A_{11}, A_{21}) \cap M^P_\wedge(C_{11}, C_{21})$, then by the IH, we can find D_{11}, D_{12}, D_{13} and D_{14} as in the statement of the lemma.

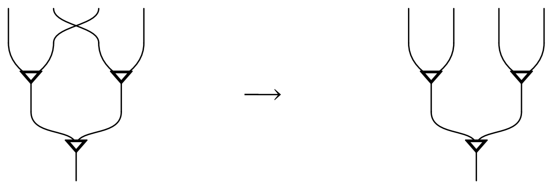
Depending on which case holds, we can now set $(B_{11}, B_{21}, B_{31}, B_{41}) \equiv (A_{111}, A_{112}, A_{211}, A_{212})$, $(C_{111}, C_{211}, C_{112}, C_{212})$ or $(D_{11}, D_{12}, D_{13}, D_{14})$. In the same way, we can find appropriate $(B_{12}, B_{22}, B_{32}, B_{42})$ from D_2. It is then straightforward to see that setting $B_i \equiv B_{i1} \; \alpha \; B_{i2}$ gives us suitable B_i.

□

Remark 2.46. The fact that atomic contractions are simply special cases of merge contractions is illustrated by the two lemmas above. The first is analogous to a rewrite of a non-atomic contraction and a contraction to three atomic contractions:

$$\mathsf{c}{\downarrow}-\mathsf{ac}{\downarrow}: \quad \mathsf{c}{\downarrow}\,\dfrac{(a \vee a) \vee (a \vee a)}{\mathsf{ac}{\downarrow}\,\dfrac{a \vee a}{a}} \quad \longrightarrow \quad \mathsf{ac}{\downarrow}\,\dfrac{\mathsf{ac}{\downarrow}\,\dfrac{a \vee a}{a} \vee \mathsf{ac}{\downarrow}\,\dfrac{a \vee a}{a}}{a}$$

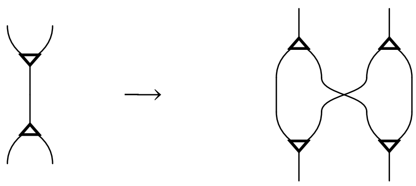
$$\text{Lemma 3.38}: \quad \mathsf{mc}{\downarrow}\,\dfrac{(A_1 \vee B_1) \vee (A_2 \vee B_2)}{\mathsf{mc}{\downarrow}\,\dfrac{A \vee B}{C}} \quad \longrightarrow \quad \mathsf{mc}{\downarrow}\,\dfrac{\mathsf{mc}{\downarrow}\,\dfrac{A_1 \vee B_1}{C_1} \vee \mathsf{mc}{\downarrow}\,\dfrac{A_2 \vee B_2}{C_2}}{C}$$



whereas the second mirrors the interaction between an atomic contraction and cocontraction.

$$\mathsf{ac}{\downarrow}-\mathsf{ac}{\uparrow}: \quad \mathsf{ac}{\downarrow}\,\dfrac{a \vee a}{\mathsf{ac}{\uparrow}\,\dfrac{a}{a \wedge a}} \quad \longrightarrow \quad \mathsf{m}\,\dfrac{\mathsf{ac}{\uparrow}\,\dfrac{a}{a \wedge a} \vee \mathsf{ac}{\uparrow}\,\dfrac{a}{a \wedge a}}{\mathsf{ac}{\downarrow}\,\dfrac{a \vee a}{a} \wedge \mathsf{ac}{\downarrow}\,\dfrac{a \vee a}{a}}$$

$$\text{Lemma 3.39}: \quad \mathsf{mc}{\downarrow}\,\dfrac{A_1 \vee A_2}{\mathsf{mc}{\uparrow}\,\dfrac{B}{C_1 \wedge C_2}} \quad \longrightarrow \quad \mathsf{m}\,\dfrac{\mathsf{mc}{\uparrow}\,\dfrac{A_1}{B_1 \wedge B_2} \vee \mathsf{mc}{\uparrow}\,\dfrac{A_2}{B_3 \wedge B_4}}{\mathsf{mc}{\downarrow}\,\dfrac{B_1 \vee B_3}{C_1} \wedge \mathsf{mc}{\downarrow}\,\dfrac{B_2 \vee B_4}{C_2}}$$



2.3.2 Permuting rules through merge contractions

We are now ready to prove a major theorem: that a merge contraction can be permuted past any other rule in SKSm. Not only does the move to SKSm greatly reduce the number of cases we need to consider, but we are now able to push medials—in particular, critical medials—past all other rules, in the guise of a merge contraction.

Theorem 2.47. *Assume we have a merge contraction above a rule $\rho \in \mathsf{SKSm}$ in a context, with P_C not a K_C-subformula of A or B:*

$$\mathsf{mc}\!\downarrow \frac{A \vee B}{K_C\left\{\rho \dfrac{P_C}{Q_C}\right\}}$$

Then we can find sentences P_A, P_B, Q_A, Q_B, contexts $K_A\{\ \}, K_B\{\ \}$ and derivations $\phi_1, \phi_2 \in \mathsf{SKSm}$ s.t. the following is a valid derivation:

$$\mathsf{mc}\!\downarrow \frac{K_A\left\{\begin{array}{c} P_A \\ \phi_1 \| \\ Q_A \end{array}\right\} \vee K_B\left\{\begin{array}{c} P_B \\ \phi_2 \| \\ Q_B \end{array}\right\}}{K_C\{Q_C\}}$$

Proof. By Lemma 3.37, we can find K_A, K_B, C_1, C_2 so that we have:

$$\mathsf{mc}\!\downarrow \frac{K_A\{P_A\} \vee K_B\{P_B\}}{K_C\left\{\rho \dfrac{P_C}{Q_C}\right\}}$$

One general case is very straightforward:

Eq If $P_A \equiv P_B \equiv P_C$, then, by Lemma 3.37, we can set $\phi_1 \equiv \phi_2 \equiv \rho \dfrac{P_C}{Q_C}$.

We now prove case by case on ρ that we can find Q_A, Q_B, ϕ_1 and ϕ_2

- If $\rho \in \{\mathsf{ai}\!\downarrow, \mathsf{ai}\!\uparrow, \mathsf{aw}\!\downarrow, \mathsf{aw}\!\uparrow\}$, then we can observe that every possible (non-K_C-subformula) case falls under **Eq**.
- If $\rho = \mathsf{mc}\!\downarrow$, with $P_C \equiv R \vee S$ and $Q_C \equiv T$, then we have two extra cases beyond **Eq** to consider:
  - $P_A \equiv R, Q_A \equiv S$. Then, by Lemma 3.37 we have:

$$\mathsf{mc}\!\downarrow \frac{K_A\{R\} \vee K_B\{S\}}{K_C\left\{\mathsf{mc}\!\downarrow \dfrac{R \vee S}{T}\right\}} \quad \longrightarrow \quad \mathsf{mc}\!\downarrow \frac{K_A\{R\} \vee K_B\{S\}}{K_C\{T\}}$$

  - $P_A \equiv R_1 \vee S_1, Q_A \equiv R_2 \vee S_2$, with $R \in M_\vee(R_1, R_2)$ and $S \in M_\vee(S_1, S_2)$. By Lemma 3.38, we can find T_1 and T_2 s.t.:

$$\mathsf{mc}\!\downarrow \frac{K_A\{(R_1 \vee S_1)\} \vee K_B\{(R_2 \vee S_2)\}}{K_C\left\{\mathsf{mc}\!\downarrow \dfrac{R \vee S}{T}\right\}} \quad \longrightarrow \quad \mathsf{mc}\!\downarrow \frac{K_A\left\{\mathsf{mc}\!\downarrow \dfrac{R_1 \vee S_1}{T_1}\right\} \vee K_B\left\{\mathsf{mc}\!\downarrow \dfrac{R_2 \vee S_2}{T_2}\right\}}{K_C\{T\}}$$

- If $\rho = \mathsf{mc}\!\uparrow$ with $P_C \equiv T$ and $Q_C \equiv R \wedge S$, then there is one extra case beyond **Eq** to consider:

– $P_A \equiv T_A, P_B \equiv T_B$, with $T \in M_\vee(T_1, T_2)$. By Lemma 3.39, we can find R_A, R_B, S_A, S_B s.t.:

$$\mathsf{mc}\!\downarrow \frac{K_A\{T_A\} \vee K_B\{T_B\}}{K_C\left\{\mathsf{mc}\!\uparrow \frac{T}{R \wedge S}\right\}} \quad \longrightarrow \quad \mathsf{mc}\!\downarrow \frac{K_A\left\{\mathsf{mc}\!\uparrow \frac{T_A}{R_A \wedge S_A}\right\} \vee K_B\left\{\mathsf{mc}\!\uparrow \frac{T_B}{R_B \wedge S_B}\right\}}{K_C\{R \wedge S\}}$$

- If $\rho \in \{\mathsf{s}\!\downarrow, \mathsf{s}\!\uparrow\}$, then, if **Eq** does not apply, we must have a LHS of the form:

$$\mathsf{mc}\!\downarrow \frac{K_A\{(P_5 \wedge P_6)\} \vee K_B\{(P_7 \wedge P_8)\}}{K_C\left\{\mathsf{s}\!\downarrow \frac{(P_1 \;\alpha\; P_2) \wedge (P_3 \vee P_4)}{(P_1 \wedge P_3) \vee (P_2 \;\beta\; P_4)}\right\}}$$

with $(\alpha, \beta) \in \{(\vee, \wedge), (\wedge, \vee)\}$, $P_1 \;\alpha\; P_2 \in M_\vee(P_5, P_7)$ and $P_3 \vee P_4 \in M_\vee(P_6, P_8)$. Therefore, using projective derivations, we can construct:

$$\mathsf{mc}\!\downarrow \frac{K_A\left\{\mathsf{s}\!\downarrow \frac{\begin{array}{ccc} P_5 & & P_6 \\ \pi_{P_5}\| & \wedge & \pi_{P_6}\| \\ P_1 \;\alpha\; P_2 & & P_3 \vee P_4 \end{array}}{(P_1 \wedge P_3) \vee (P_2 \;\beta\; P_4)}\right\} \vee K_B\left\{\mathsf{s}\!\downarrow \frac{\begin{array}{ccc} P_7 & & P_8 \\ \pi_{P_7}\| & \wedge & \pi_{P_8}\| \\ P_1 \;\alpha\; P_2 & & P_3 \vee P_4 \end{array}}{(P_1 \wedge P_3) \vee (P_2 \;\beta\; P_4)}\right\}}{K_C\{(P_1 \wedge P_3) \vee (P_2 \;\beta\; P_4)\}}$$

□

Definition 2.48. In the following definition, $\alpha, \beta, \delta, \epsilon \in \{\vee, \wedge\}$.

We define the following rewrite rules for SKSm:

$$r_1 : \; \rho \frac{\mathsf{mc}\!\downarrow \frac{(A_1 \;\alpha\; A_2) \vee (B_1 \;\alpha\; B_2)}{C_1 \;\alpha\; C_2} \wedge (D \;\beta\; E)}{(C_1 \wedge D) \;\delta\; (C_2 \;\epsilon\; E)} \quad \longrightarrow \quad \mathsf{mc}\!\downarrow \frac{\mathsf{s}\!\uparrow \frac{((A_1 \;\alpha\; A_2) \vee (B_1 \;\alpha\; B_2)) \wedge \mathsf{mc}\!\uparrow \frac{(D \;\beta\; E)}{(D \;\beta\; E) \wedge (D \;\beta\; E)}}{\rho \frac{(A_1 \;\alpha\; A_2) \wedge (D \;\beta\; E)}{(A_1 \wedge D) \;\delta\; (A_2 \;\epsilon\; E)} \vee \rho \frac{(B_1 \;\alpha\; B_2) \wedge (D \;\beta\; E)}{(B_1 \wedge D) \;\delta\; (B_2 \;\epsilon\; E)}}}{(C_1 \wedge D) \;\delta\; (C_2 \;\epsilon\; E)}$$

where ρ is an instance of $\wedge\mathsf{c}\!\uparrow$, $\mathsf{s}\!\downarrow$ or $\mathsf{s}\!\uparrow$.

$$r_2 : \; \rho \frac{\mathsf{mc}\!\downarrow \frac{(A_1 \;\alpha\; A_2) \vee (B_1 \;\alpha\; B_2)}{C_1 \;\alpha\; C_2} \vee (D \;\beta\; E)}{(C_1 \vee D) \;\delta\; (C_2 \vee E)} \quad \longrightarrow \quad \mathsf{mc}\!\downarrow \frac{\vee\mathsf{c}\!\downarrow \frac{((A_1 \;\alpha\; A_2) \vee (B_1 \;\alpha\; B_2)) \vee \left((D \;\beta\; E) \vee \begin{array}{c} \mathsf{f} \\ \|\{\mathsf{aw}\!\downarrow\} \\ (D \;\beta\; E) \end{array}\right)}{\rho \frac{(A_1 \;\alpha\; A_2) \vee (D \;\beta\; E)}{(A_1 \vee D) \;\delta\; (A_2 \vee E)} \vee \rho \frac{(B_1 \;\alpha\; B_2) \vee (D \;\beta\; E)}{(B_1 \vee D) \;\delta\; (B_2 \vee E)}}}{(C_1 \vee D) \;\delta\; (C_2 \vee E)}$$

where ρ is an instance of m or $\vee\mathsf{c}\!\downarrow$.

$$s : \; \mathsf{mc}\!\downarrow \frac{A \vee B}{K_C\left\{\rho \frac{P_C}{Q_C}\right\}} \quad \longrightarrow \quad \mathsf{mc}\!\downarrow \frac{K_A\left\{\begin{array}{c} P_A \\ \phi_1\| \\ Q_A \end{array}\right\} \vee K_B\left\{\begin{array}{c} P_B \\ \phi_2\| \\ Q_B \end{array}\right\}}{K_C\{Q_C\}}$$

where ρ is an instance of any rule, P_C is not a K_C-subformula of A or B and the RHS is given as in Theorem 2.47.

We also define the trivial reductions:

$$t_1 : \quad \rho\,\dfrac{K\left\{\mathsf{mc}{\downarrow}\,\dfrac{A \vee B}{C}\right\}}{K'\{C\}} \quad \longrightarrow \quad \rho\,\dfrac{K\{A \vee B\}}{K'\left\{\mathsf{mc}{\downarrow}\,\dfrac{A \vee B}{C}\right\}}$$

$$t_2 : \quad \mathsf{mc}{\downarrow}\,\dfrac{A \vee B}{K_C\left\{\rho\,\dfrac{P_C}{Q_C}\right\}} \quad \longrightarrow \quad \mathsf{mc}{\downarrow}\,\dfrac{K_A\left\{\rho\,\dfrac{P_C}{Q_C}\right\} \vee B}{K_C\{Q_C\}}$$

Where ρ is an instance of any rule, and P_C is a K_C-subformula of A.

$$t_3 : \quad \mathsf{mc}{\downarrow}\,\dfrac{A \vee B}{K_C\left\{\rho\,\dfrac{P_C}{Q_C}\right\}} \quad \longrightarrow \quad \mathsf{mc}{\downarrow}\,\dfrac{A \vee K_B\left\{\rho\,\dfrac{P_C}{Q_C}\right\}}{K_C\{Q_C\}}$$

Where ρ is an instance of any rule, and P_C is a K_C-subformula of B.

We define the rewriting system $\mathrm{M} = \{r_1, r_2, s, t_1, t_2, t_3\}$.

2.4 Cycle Removal with Merge Contractions

We will now show we can use the rewriting system M to remove cycles from a proof. To prove termination we will focus on removing the critical merge medials, rather than cycles themselves. We do so by defining the image of critical medials in SKSm.

Definition 2.49. Let ϕ be an SKS proof with a critical medial m. We call the image of m in $\nu_{\min} : \mathsf{SKS} \to \mathsf{SKSm}$ or $\nu_{\max} : \mathsf{SKS} \to \mathsf{SKSm}$ a *critical merge contraction*.

By removing the critical merge contractions in an SKSm proof one at a time, we can afford to create new cycles, as long as they do not create any new critical merge contractions, since, if all critical merge contractions are eliminated, there can be no more cycles.

Given an SKSm derivation, there can be several merge contractions to which we can apply the reduction rules of system M. However, since our goal is to permute critical medials downwards until there are no cycles left in the derivation there is a clear strategy to choose which merge contractions will be permuted.

Proposition 2.50. *Let ϕ be an* SKS *derivation with a unicycle. Let $\phi' \equiv \nu_{\max}(\phi)$ (or $\nu_{\min}(\phi)$, it doesn't matter). Then ϕ' has a critical merge contraction to which a reduction rule of system M can be applied.*

Proof. Given a SKS derivation ϕ with a unicycle, there must be a critical medial. Since μ is the left inverse of $\nu_{\max}$, the merge contraction that the medial is mapped to must be a critical merge contraction.

We can apply a reduction of rule of system M to it since there is at least one rule below it: the cut belonging to the unicycle. □

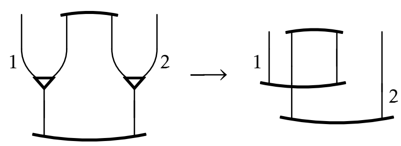
2.4.1 Termination

Our strategy for removing ai-cycles is as follows:

1. Ensure that all cycles are unicycles, and that no two cycles share a cut, by Proposition 2.20. Therefore every cycle will have a well-defined critical medial.

2. Convert the proof to SKSm, and permute the critical merge contractions down using the rewrite system M.

3. When the critical merge contraction for a cycle is permuted past the (unique) atomic cut that is a node in that cycle, in application of rewrite s, it will break the cycle. To ensure that this rewriting does indeed break the cycle and doesn't simply create a a 2-multicycle, we need to make sure that edges 1 and 2 are not connected by an atomic identity, however since no two cycles share an atomic cut, this cannot be the case.

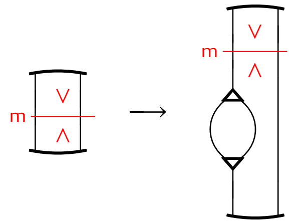


We will show termination of the procedure by proving we can remove critical merge contractions one by one, assuming all cycles are unicycles and no two cycles share a cut. Since termination can be understood from the changes induced on the flow of the derivation by the cycle-elimination procedure, we will accompany the proof of the following Theorem 2.52 with a study of the changes on flows that each rewriting rule of system M produces. An accurate formal bound for the complexity cost of the procedure has yet to be established, but the study of the flow changes is expected to provide us with the necessary intuition to obtain it.

Lemma 2.51. *If the edge of a cycle is bifurcated by rewrite r_1 then, although new cycles are created, no new critical merge contractions are created, and the two conditions in Proposition 2.20 are preserved.*

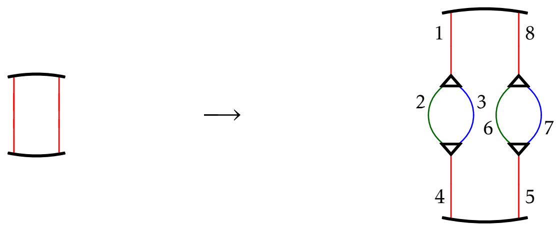
Proof. We need to consider two cases: where one edge of a cycle is bifurcated and where both are.

- If only one edge of a cycle is bifurcated, since the critical merge contraction for the cycle is above the bifurcation, it is the critical merge contraction for both new cycles as well.



- If two edges of the cycle are bifurcated, we must be in the following situation:

$$
\begin{array}{c}
P\left\{\dfrac{\mathsf{t}}{a \vee \bar{a}}\right\} \\
\| \\
S\left\{ \rho\,\dfrac{\boxed{M\,\dfrac{A \vee B}{C}} \wedge K\{a\}\{\bar{a}\}}{F\{a\}\{\bar{a}\}} \right\} \\
\| \\
T\left\{\dfrac{a \wedge \bar{a}}{\mathsf{f}}\right\}
\end{array}
\longrightarrow
\begin{array}{c}
P\left\{\dfrac{\mathsf{t}}{a \vee \bar{a}}\right\} \\
\| \\
S\left\{ M\,\dfrac{\mathsf{s}\!\uparrow\dfrac{(A \vee B) \wedge \boxed{\mathsf{mc}\!\uparrow\dfrac{K\{a\}\{\bar{a}\}}{K\{a\}\{\bar{a}\} \wedge K\{a\}\{\bar{a}\}}}}{\boxed{\rho\,\dfrac{A \wedge K\{a\}\{\bar{a}\}}{F\{a\}\{\bar{a}\}}} \vee \boxed{\rho\,\dfrac{B \wedge K\{a\}\{\bar{a}\}}{F\{a\}\{\bar{a}\}}}}}{F\{a\}\{\bar{a}\}} \right\} \\
\| \\
T\left\{\dfrac{a \wedge \bar{a}}{\mathsf{f}}\right\}
\end{array}
$$



In this case, the critical merge contraction for the original cycle (on the left) is now the critical merge contraction for the cycles $(1,2,4,5,6,8)$ and $(1,3,4,5,7,8)$ where a and $\bar{a}$ are of the same color. M is the critical merge contraction for the cycles $(1,2,4,5,7,8)$ and $(1,3,4,5,6,8)$ where a and $\bar{a}$ are of different colors. Thus, although we do add cycles, we do not add critical merge contractions.

$\square$

We now have all the pieces in play to prove our main theorem.

Theorem 2.52. *Let ϕ be an* SKSm *derivation with no multicycles and where no two cycles with different identities share a cut. If ϕ contains n critical merge contractions, then exists a derivation ψ with the same premise and conclusion with*

$n-1$ critical merge contractions. Furthermore, ψ contains no multicycles and no two cycles with different identities share a cut.

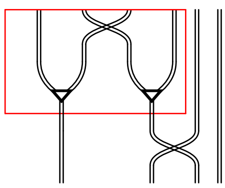
Proof. We consider the lowest critical merge contraction of ϕ, that we call M. We apply a reduction of system M to permute M downwards in ϕ. We repeat this process until we obtain a derivation ψ where M is not a critical merge contraction as it has been permuted below the cuts of the cycles whose critical medials it contained.

At every application of a reduction of M, the number of inference rules below M is decreased: the procedure therefore terminates, and at the end of it M will no longer be critical. We only need to show that we do not create multicycles, cycles with different identities that share a cut, or new critical merge contractions. We will show this by considering ten different cases, most of which will be illustrated with reference to their flows.

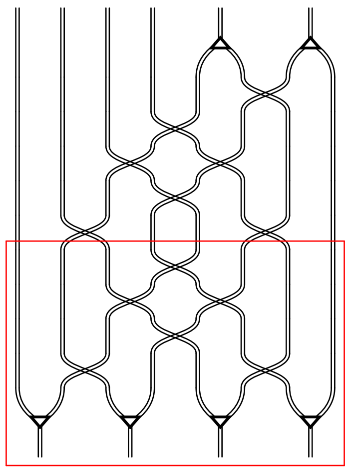
We will enclose the parts of the flow that belong to the critical merge contraction in a *red* box and we will call ρ the rule instance below the critical merge contraction.

1. Instances of r_1 do not change the links between the existing edges of a flow. They may bifurcate previously "single" edges, in which case we know from Lemma 2.51 that no new critical merge contractions are created.

$$\rho\,\frac{\mathsf{mc}{\downarrow}\,\dfrac{(A_1\ \alpha\ A_2)\vee(B_1\ \alpha\ B_2)}{C_1\ \alpha\ C_2}\wedge(D\ \beta\ E)}{(C_1\wedge D)\ \delta\ (C_2\ \epsilon\ E)} \quad\longrightarrow\quad \mathsf{mc}{\downarrow}\,\frac{\mathsf{s}{\uparrow}\,\dfrac{((A_1\ \alpha\ A_2)\vee(B_1\ \alpha\ B_2))\wedge\mathsf{mc}{\uparrow}\,\dfrac{(D\ \beta\ E)}{(D\ \beta\ E)\wedge(D\ \beta\ E)}}{\rho\,\dfrac{(A_1\ \alpha\ A_2)\wedge(D\ \beta\ E)}{(A_1\wedge D)\ \delta\ (A_2\ \epsilon\ E)}\vee\rho\,\dfrac{(B_1\ \alpha\ B_2)\wedge(D\ \beta\ E)}{(B_1\wedge D)\ \delta\ (B_2\ \epsilon\ E)}}}{(C_1\wedge D)\ \delta\ (C_2\ \epsilon\ E)}$$



$\longrightarrow$

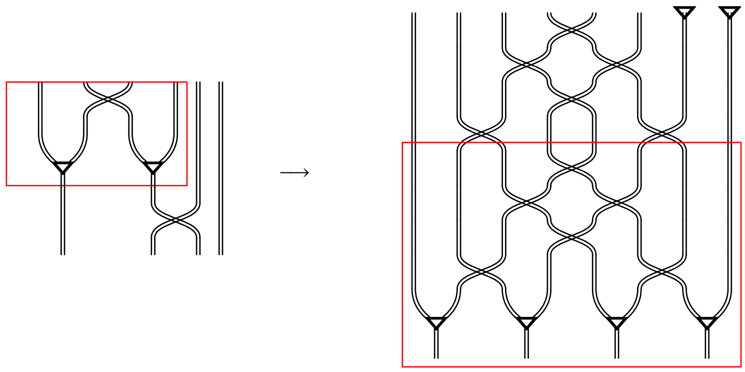


Potential complexity is generated in the cycle-removing procedure by turning straight edges into 'sausages'

2. Instances of r_2 do not change the links between the existing edges of a flow. They introduce some contractions where one edge is connected to a

weakening. Therefore the application of this rule cannot create new cycles (i.e. create new critical merge contractions) or change the identities or cuts of cycles.

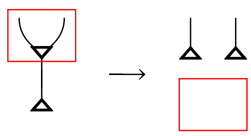
$$
\rho\,\frac{\mathsf{mc}{\downarrow}\,\dfrac{(A_1\;\alpha\;A_2)\vee(B_1\;\alpha\;B_2)}{C_1\;\alpha\;C_2}\vee(D\;\beta\;E)}{(C_1\vee D)\;\delta\;(C_2\vee E)}
\quad\longrightarrow\quad
\mathsf{mc}{\downarrow}\,\frac{\mathsf{mc}{\downarrow}\,\dfrac{((A_1\;\alpha\;A_2)\vee(B_1\;\alpha\;B_2))\vee\left((D\;\beta\;E)\vee\begin{array}{c}\mathsf{f}\\ \Vert\{\mathsf{aw}{\downarrow}\}\\ (D\;\beta\;E)\end{array}\right)}{\rho\,\dfrac{(A_1\;\alpha\;A_2)\vee(D\;\beta\;E)}{(A_1\vee D)\;\delta\;(A_2\vee E)}\vee\rho\,\dfrac{(B_1\;\alpha\;B_2)\vee(D\;\beta\;E)}{(B_1\vee D)\;\delta\;(B_2\vee E)}}}{(C_1\vee D)\;\delta\;(C_2\vee E)}
$$



The size of the proof after the cycle elimination will not be increased significantly by the application of these reductions, since the weakenings can be pulled down, and the edges that have been connected to a weakening node will return to simply being straight edges.

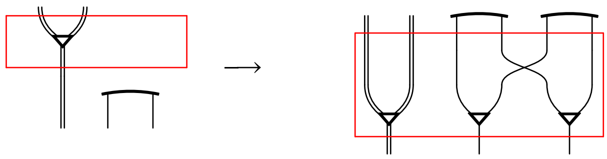
3. For all instances of s where $\rho = \mathsf{aw}{\downarrow}$ or $\rho = \mathsf{aw}{\uparrow}$, **Eq** applies. This duplicates (co)weakenings, with no risk of affecting cycles.

$$
\mathsf{mc}{\downarrow}\,\frac{A\vee B}{C\left\{\mathsf{aw}{\uparrow}\,\dfrac{a}{\mathsf{t}}\right\}}
\quad\longrightarrow\quad
\mathsf{mc}{\downarrow}\,\frac{\begin{array}{c}A\\ \pi_A\,\Vert\{\mathsf{aw}{\downarrow}\}\\ C\left\{\mathsf{aw}{\uparrow}\,\dfrac{a}{\mathsf{t}}\right\}\end{array}\vee\begin{array}{c}B\\ \pi_B\,\Vert\{\mathsf{aw}{\downarrow}\}\\ C\left\{\mathsf{aw}{\uparrow}\,\dfrac{a}{\mathsf{t}}\right\}\end{array}}{C\{\mathsf{t}\}}
$$



4. For all instances of s where $\rho = \mathsf{ai}{\downarrow}$, **Eq** applies. This changes a single identity into two identities and two contractions, and may introduce some contractions where one edge is connected to a weakening.

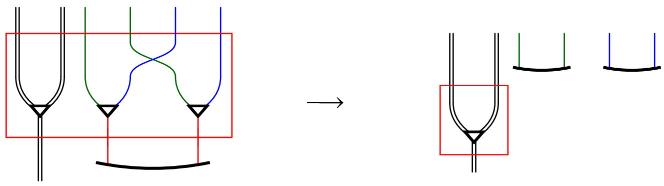
$$\mathsf{mc}{\downarrow}\,\frac{K_A\{\mathsf{t}\} \vee K_B\{\mathsf{t}\}}{K_C\left\{\mathsf{ai}{\downarrow}\,\dfrac{\mathsf{t}}{a \vee \bar{a}}\right\}} \quad\longrightarrow\quad \mathsf{mc}{\downarrow}\,\frac{K_A\left\{\mathsf{ai}{\downarrow}\,\dfrac{\mathsf{t}}{a \vee \bar{a}}\right\} \vee K_B\left\{\mathsf{ai}{\downarrow}\,\dfrac{\mathsf{t}}{a \vee \bar{a}}\right\}}{K_C\{a \vee \bar{a}\}}$$



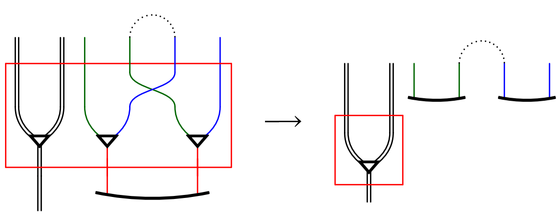
This reduction could only introduce a cycle if the instance of ai↓ being permuted was part of a cycle. This is however not a possible case since we are permuting the lowermost critical merge contraction, and if the instance of ai↓ was part of a cycle there would be a critical merge contraction that is lower.

5. For all instances of s where $\rho = \mathsf{ai}{\uparrow}$, **Eq** applies. This duplicates cuts and removes atomic contractions:

$$\mathsf{mc}{\downarrow}\,\frac{K_A\{(a \wedge \bar{a})\} \vee K_B\{(a \wedge \bar{a})\}}{K_C\left\{\mathsf{i}{\uparrow}\,\dfrac{a \wedge \bar{a}}{\mathsf{f}}\right\}} \quad\longrightarrow\quad \mathsf{mc}{\downarrow}\,\frac{K_A\left\{\mathsf{i}{\uparrow}\,\dfrac{a \wedge \bar{a}}{\mathsf{f}}\right\} \vee K_B\left\{\mathsf{i}{\uparrow}\,\dfrac{a \wedge \bar{a}}{\mathsf{f}}\right\}}{K_C\{\mathsf{f}\}}$$



If a and $\bar{a}$ are of the same color are in a cycle, then the critical merge contraction for the cycle is above the subderivation that we are rewriting, and remains the critical medial for the cycle. If a and $\bar{a}$ of different colors belong to a cycle, then M is its critical merge contraction, and the cycle is broken through this reduction: since there are no cycles with different identities sharing a cut, the other two edges cannot be connected.



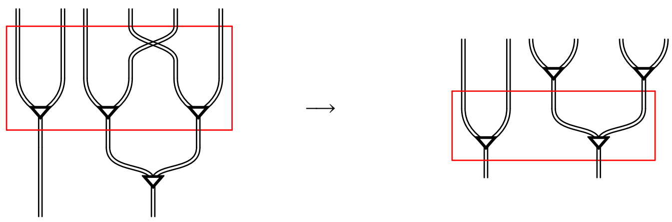
Therefore the application of this rule reduces or maintains the number of critical merge contractions), and does not create multicycles or cycles with different identities that share a cut.

6. Applications of s where $\rho = \mathsf{mc}{\downarrow}$ do not change the flow significantly in any of the three cases:

   - In the **Eq** case, and in the second of the two extra cases, the critical merge contraction is permuted beneath the other, duplicating it but making no major change to the flow.

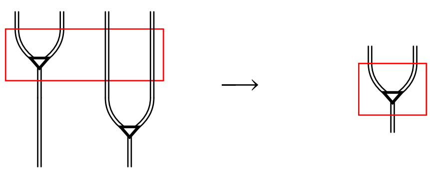
$$\mathsf{mc}{\downarrow}\,\frac{K_A\{(R\vee S)\}\vee K_B\{(R\vee S)\}}{K_C\left\{\mathsf{mc}{\downarrow}\,\dfrac{R\vee S}{T}\right\}} \longrightarrow \mathsf{mc}{\downarrow}\,\frac{K_A\left\{\mathsf{mc}{\downarrow}\,\dfrac{R\vee S}{T}\right\}\vee K_B\left\{\mathsf{mc}{\downarrow}\,\dfrac{R\vee S}{T}\right\}}{K_C\{T\}}$$

$$\mathsf{mc}{\downarrow}\,\frac{K_A\{(R_1\vee S_1)\}\vee K_B\{(R_2\vee S_2)\}}{K_C\left\{\mathsf{mc}{\downarrow}\,\dfrac{R\vee S}{T}\right\}} \longrightarrow \mathsf{mc}{\downarrow}\,\frac{K_A\left\{\mathsf{mc}{\downarrow}\,\dfrac{R_1\vee S_1}{T_1}\right\}\vee K_B\left\{\mathsf{mc}{\downarrow}\,\dfrac{R_2\vee S_2}{T_2}\right\}}{K_C\{T\}}$$



   - In the first additional case, the merge contraction below simply gets absorbed by the critical merge contraction.

$$\mathsf{mc}{\downarrow}\,\frac{K_A\{R\}\vee K_B\{S\}}{K_C\left\{\mathsf{mc}{\downarrow}\,\dfrac{R\vee S}{T}\right\}} \longrightarrow \mathsf{mc}{\downarrow}\,\frac{K_A\{R\}\vee K_B\{S\}}{K_C\{T\}}$$

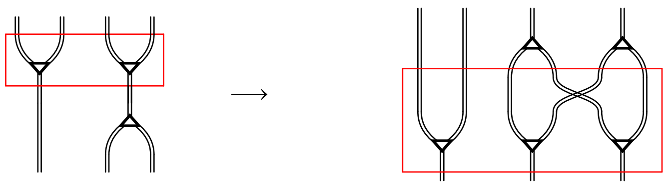


7. Applications of s where $\rho = \mathsf{mc}{\uparrow}$ change the flow a little, but do not change the paths through the sub flow. Both the cases can be treated

together:

$$\mathsf{mc}\!\downarrow \frac{K_A\{T\} \vee K_B\{T\}}{K_C\left\{\mathsf{mc}\!\downarrow \dfrac{T}{R \vee S}\right\}} \quad \longrightarrow \quad \mathsf{mc}\!\downarrow \frac{K_A\left\{\mathsf{mc}\!\uparrow \dfrac{T}{R \wedge S}\right\} \vee K_B\left\{\mathsf{mc}\!\uparrow \dfrac{T}{R \wedge S}\right\}}{K_C\{R \wedge S\}}$$

$$\mathsf{mc}\!\downarrow \frac{K_A\{T_A\} \vee K_B\{T_B\}}{K_C\left\{\mathsf{mc}\!\downarrow \dfrac{T}{R \vee S}\right\}} \quad \longrightarrow \quad \mathsf{mc}\!\downarrow \frac{K_A\left\{\mathsf{mc}\!\uparrow \dfrac{T_A}{R_A \wedge S_A}\right\} \vee K_B\left\{\mathsf{mc}\!\uparrow \dfrac{T_B}{R_B \wedge S_B}\right\}}{K_C\{R \wedge S\}}$$

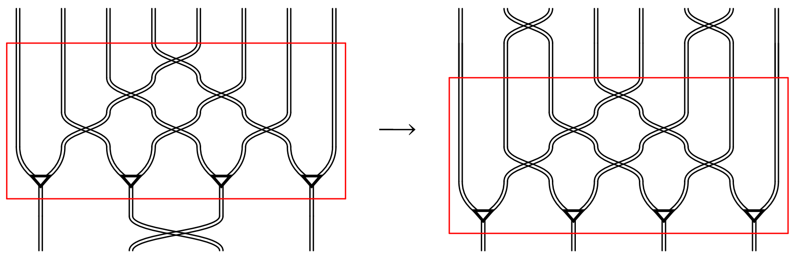


8. Instances where $\rho \in \{\mathsf{s}\!\downarrow, \mathsf{s}\!\uparrow\}$ can be treated together, as the flows are indistinguishable.

   - If **Eq** holds, the flow changes but all the paths are the same:

$$\mathsf{mc}\!\downarrow \frac{K_A\{((P_1 \;\alpha\; P_2) \wedge (P_3 \vee P_4))\} \vee K_B\{((P_1 \;\alpha\; P_2) \wedge (P_3 \vee P_4))\}}{K_C\left\{\mathsf{s}\!\downarrow \dfrac{(P_1 \;\alpha\; P_2) \wedge (P_3 \vee P_4)}{(P_1 \wedge P_3) \vee (P_2 \;\beta\; P_4)}\right\}}$$

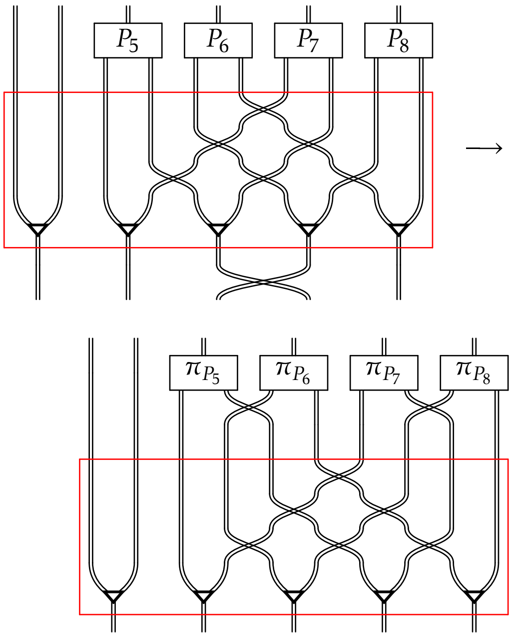
$$\longrightarrow \quad \mathsf{mc}\!\downarrow \frac{K_A\left\{\mathsf{s}\!\downarrow \dfrac{(P_1 \;\alpha\; P_2) \wedge (P_3 \vee P_4)}{(P_1 \wedge P_3) \vee (P_2 \;\beta\; P_4)}\right\} \vee K_B\left\{\mathsf{s}\!\downarrow \dfrac{(P_1 \;\alpha\; P_2) \wedge (P_3 \vee P_4)}{(P_1 \wedge P_3) \vee (P_2 \vee P_4)}\right\}}{K_C\{(P_1 \wedge P_3) \vee (P_2 \;\beta\; P_4)\}}$$



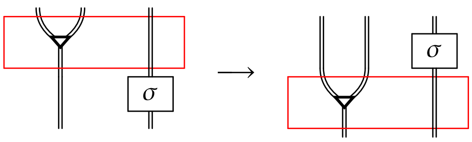
   - If not, then we have the rewriting involving projective derivations, which, since it does not introduce any new paths, cannot cause problems:

$$\mathsf{mc}\!\downarrow \frac{K_A\{(P_5 \wedge P_6)\} \vee K_B\{(P_7 \wedge P_8)\}}{K_C\left\{\mathsf{s}\!\downarrow \dfrac{(P_1 \;\alpha\; P_2) \wedge (P_3 \vee P_4)}{(P_1 \wedge P_3) \vee (P_2 \;\beta\; P_4)}\right\}} \longrightarrow$$

$$\mathsf{mc}\!\downarrow \dfrac{K_A\left\{\mathsf{s}\!\downarrow \dfrac{\begin{array}{ccc} P_5 & & P_6 \\ \pi_{P_5}\,\| & \wedge & \pi_{P_6}\,\| \\ P_1 \;\alpha\; P_2 & & P_3 \vee P_4 \end{array}}{(P_1 \wedge P_3) \vee (P_2 \;\beta\; P_4)}\right\} \vee K_B\left\{\mathsf{s}\!\downarrow \dfrac{\begin{array}{ccc} P_7 & & P_8 \\ \pi_{P_7}\,\| & \wedge & \pi_{P_8}\,\| \\ P_1 \;\alpha\; P_2 & & P_3 \vee P_4 \end{array}}{(P_1 \wedge P_3) \vee (P_2 \;\beta\; P_4)}\right\}}{K_C\,\{(P_1 \wedge P_3) \vee (P_2 \;\beta\; P_4)\}}$$



9. Instances of the trivial reductions t_1, t_2 or t_3 do not change the flow of the derivation and therefore cannot produce new cycles.



□

To eliminate all cycles from a derivation, one simply performs the procedure n times, once for each critical merge contraction.

Theorem 2.53. *Given an* SKS *derivation ϕ, we can construct an* SKS *derivation ψ with the same premise and conclusion containing no* ai*-cycles.*

Proof. Given a derivation ϕ, we transform all its multicycles into unicycles by applying Lemma 2.18. If any two cycles share a cut, we apply Lemma 2.19. Call this proof ϕ'. Take $\chi \equiv \nu_{max}(\phi')$. We can eliminate every critical merge contraction from χ with an application of Theorem 2.52, obtaining χ'. Finally we take $\psi \equiv \mu(\chi')$, which will be free of cycles. □

A complete example of the removal of a cycle with a similar procedure can be found in [AT16].

We can now prove a full decomposition result.

Theorem 2.54. *Given an* SKS *derivation* ϕ *of* A *to* B*, we can obtain a strongly decomposed derivation* ϕ *from* A *to* B*.*

Proof. Immediate from Theorems 2.53 and 2.15. □

Note that we could define a variation of decomposition for proofs in SKSm itself, if we add a rewrite to "minimize" merge contractions, so that any rewrite in C can be replicated in SKSm.

Definition 2.55. We define the following rewrite for derivations in SKSm,

$$m : \phi \longrightarrow \nu_{\min}(\mu(\phi))$$

In addition, we define $\mathsf{mc}\!\downarrow - \mathsf{ai}\!\uparrow$ and $\mathsf{mc}\!\uparrow - \mathsf{ai}\!\downarrow$ to be identical to $\mathsf{ac}\!\downarrow - \mathsf{ai}\!\uparrow$ and $\mathsf{ac}\!\uparrow - \mathsf{ai}\!\downarrow$ respectively, except with different labels for the contractions. We define $\mathrm{M}^+ = \mathrm{M} \cup \{m, \mathsf{mc}\!\downarrow - \mathsf{ai}\!\uparrow, \mathsf{mc}\!\uparrow - \mathsf{ai}\!\downarrow\}$.

Proposition 2.56. *If* $\phi \longrightarrow_{\mathrm{C}} \psi$*, then* $\nu_{\min}(\phi) \longrightarrow_{\mathrm{M}^+} \nu_{\min}(\psi)$*.*

Theorem 2.57. *For every proof* ϕ *in* SKSm*, there is a proof* ψ *and a reduction sequence* $\phi \longrightarrow^*_{\mathrm{M}^+} \psi$*, such that*

$$\nu(\psi) \equiv \begin{array}{l} \| \mathsf{SKS}\setminus\{\mathsf{ac}\!\downarrow,\mathsf{ac}\!\uparrow\} \\ B' \\ \| \{\mathsf{ac}\!\downarrow\} \\ B \end{array}$$

Proof. We follow the strategy of "do cycle removal then do decomposition". □

It is expected that a stronger result is true for derivations in SKSm.

Conjecture 2.58. M^+ *is weakly normalizing for* SKSm*. Therefore every derivation can be reduced to the following form:*

$$\begin{array}{l} A \\ \| \{\mathsf{mc}\!\uparrow\} \\ A' \\ \| \mathsf{SKSm}\setminus\{\mathsf{mc}\!\downarrow,\mathsf{mc}\!\uparrow\} \\ B' \\ \| \{\mathsf{mc}\!\downarrow\} \\ B \end{array}$$

Part II

First-Order Logic

Chapter 3

Open Deduction for First-Order Logic

3.1 Open Deduction and First-Order Logic

3.1.1 The signature for classical first-order logic

Definition 3.1. We define $\mathcal{V} = \{x_0, x_1, \dots\}$ and $\mathcal{C} = \{c, c_0, c_1, \dots\}$ to be infinite sets of variables and constants, respectively. For each n, we define $\mathcal{F}_n = \{f_0^n, f_1^n, \dots\}$ to be an infinite set of n-ary function symbols.

We define the *terms* of FOL, $\mathcal{T}$, in the following way:

- All variables $x \in \mathcal{V}$ are terms.
- All constants $a \in \mathcal{C}$ are terms.
- If $t_1, \dots, t_n$ are all terms, and $f_n \in \mathcal{F}_n$, then $\mathcal{F}_n(t_1, \dots, t_n)$ is a term.

A term is *closed* if it contains no variables.

Definition 3.2. For each n, we define $\mathcal{P}_n = \{P_0^n, P_1^n, \dots\}$ to be an infinite set of *n-ary predicate symbols*. We define the set of atoms of first-order logic to be $\mathcal{A}_1 = \{P_i^n(t_1, \dots, t_n) \mid i, n \in \mathbb{N}, t_i \in \mathcal{T}\}$.

We define $\mathcal{Q} = \{\forall x_0, \forall x_1, \dots\} \cup \{\exists x_0, \exists x_1, \dots\}$, to be an infinite set of logical operators of arity 1, s.t. for every $v \in \mathcal{V}$, $\forall v, \exists v \in \mathcal{Q}$, with $\bar{\forall} v = \exists v$ and $\bar{\exists} v = \forall v$

Definition 3.3. We define the signature for FOL, $\Sigma_1 = (\mathcal{A}_1, \{\mathsf{t}, \mathsf{f}\}, \mathcal{Q}, \{\wedge, \vee\})$.

Convention 3.4. In practice, we will use x, y, z for variables, a, b, c for constants, f, g, h for functions of any arity and P, Q, R for predicate symbols of any arity, all with or without subscripts and superscripts.

Definition 3.5. The size of a term, denoted $|t|$, is defined as follows:

- $|x| = 1$ for every $x \in \mathcal{V}$
- $|a| = 1$ for every $a \in \mathcal{C}$

- $f_n(t_1,\ldots,t_n) = \max(t_1,\ldots,t_n) + 1$ for every $f_n \in \mathcal{F}_n, t_i \in \mathcal{T}$.

The set of *free variables* of a term, $FV(t)$ is defined as follows:

- $FV(x) = \{x\}$,
- $FV(a) = \emptyset$,
- $FV(f_n(t_1,\ldots,t_n)) = \bigcup(FV(t_i))$ for every $f_n \in \mathcal{F}_n, t_i \in \mathcal{T}$.

If $t \in \mathcal{T}$ and $x \in \mathcal{V}$, then $[t \Rightarrow x]$ denotes a *substitution*. If $t' \in \mathcal{T}$, then we write $[t \Rightarrow x]t'$ to denote the term defined in the following way:

- $[t \Rightarrow x]x = t$,
- If $x \neq y$, $[t \Rightarrow x]y = y$.
- $[t \Rightarrow x]a = t$,
- $[t \Rightarrow x](f_n(t_1,\ldots,t_n)) = f_n([t \Rightarrow x]t_1,\ldots,[t \Rightarrow x]t_n)$

Definition 3.6. The size of a FOL derivation is determined by the following functions:

- $|\mathsf{t}| = |\mathsf{f}| = 1$,
- $\left|P_i^n(t_1,\ldots,t_n)\right| = \left(\Sigma_1^n |t_i|\right) + 1$,

The set of free variables of a FOL derivation is defined in the following way:

- $FV(\mathsf{t}) = FV(\mathsf{f}) = \emptyset$,
- $FV(P_i^n(t_1,\ldots,t_n)) = \bigcup(FV(t_i))$,
- $FV(\forall v\phi) = FV(\exists v\phi) = FV(\phi)\backslash\{v\}$,
- $FV(\phi \wedge \psi) = FV(\phi \vee \psi) = FV\left(\rho\,\dfrac{\phi}{\psi}\right) = FV(\phi) \cup FV(\psi)$.

We say that a FOL derivation ϕ is *closed* if $FV(\phi) = \emptyset$.

We write $[t \Rightarrow x]\phi$ to denote the derivation defined in the following way:

- $[t \Rightarrow x]\mathsf{t} \equiv \mathsf{t}, [t \Rightarrow x]\mathsf{f} \equiv \mathsf{f}$,
- $[t \Rightarrow x]P_i^n(t_1,\ldots,t_n) \equiv P_i^n([t \Rightarrow x]t_1,\ldots,[t \Rightarrow x]t_n)$,
- $[t \Rightarrow x](\forall x\phi) \equiv \forall x\phi, [t \Rightarrow x](\exists x\phi) \equiv \exists x\phi$
- If $x \neq y$, $[t \Rightarrow x](\forall x\phi) \equiv \forall x([t \Rightarrow x]\phi), [t \Rightarrow x](\exists x\phi) \equiv \exists x([t \Rightarrow x]\phi)$
- $[t \Rightarrow x](\phi \wedge \psi) \equiv [t \Rightarrow x]\phi \wedge [t \Rightarrow x]\psi, [t \Rightarrow x](\phi \vee \psi) \equiv [t \Rightarrow x]\phi \vee [t \Rightarrow x]\psi$
- $[t \Rightarrow x]\left\{\rho\,\dfrac{\phi}{\psi}\right\} = \rho\,\dfrac{[t \Rightarrow x]\phi}{[t \Rightarrow x]\psi}$

Convention 3.7. If $K\{\ \}$ is a first-order context and A a first-order formula, $[t \Rightarrow x]K\{A\}$ only denotes a substitution on $K\{\ \}$, $[t \Rightarrow x](K\{A\})$ denotes a substitution on all of $K\{A\}$. We will also abuse notation a little sometimes, writing, for example, $[x \Rightarrow c]A$ for A' if $[c \Rightarrow x]A' \equiv A$ and c does not appear in A'.

3.1.2 Proofs systems for FOL: SKSgq, SKSq, KSgq and KSq

We are now ready to introduce the basic open deduction proof systems for FOL, as defined in [Brü03a; Brü06b]. As for propositional logic, there are two systems with atomic structural rules, and two with general structural rules; there are two systems with cut, and two cut-free systems.

Definition 3.8. We define four proof systems for FOL,

- SKSgq, a general first-order system with cut;

$$\mathsf{SKSgq} = \mathsf{SKSg} + \boxed{\begin{array}{cc} \mathsf{n}{\downarrow}\dfrac{[t \Rightarrow x]A}{\exists x A} & \mathsf{u}{\downarrow}\dfrac{\forall x(A \vee B)}{\forall x A \vee \exists x B} \\ \\ \mathsf{n}{\uparrow}\dfrac{\forall x A}{[t \Rightarrow x]A} & \mathsf{u}{\uparrow}\dfrac{\exists x A \wedge \forall x B}{\exists x(A \wedge B)} \end{array}}$$

- SKSq, an atomic first-order system with cut;

$$\mathsf{SKSq} = \mathsf{SKS} + \boxed{\begin{array}{cccc} \mathsf{n}{\downarrow}\dfrac{[t \Rightarrow x]A}{\exists x A} & \mathsf{u}{\downarrow}\dfrac{\forall x(A \vee B)}{\forall x A \vee \exists x B} & \mathsf{m}_1{\downarrow}\dfrac{\exists x A \vee \exists x B}{\exists x(A \vee B)} & \mathsf{m}_2{\downarrow}\dfrac{\forall x A \vee \forall x B}{\forall x(A \vee B)} \\ \\ \mathsf{n}{\uparrow}\dfrac{\forall x A}{[t \Rightarrow x]A} & \mathsf{u}{\uparrow}\dfrac{\exists x A \wedge \forall x B}{\exists x(A \wedge B)} & \mathsf{m}_1{\uparrow}\dfrac{\forall x(A \wedge B)}{\forall x A \wedge \forall x B} & \mathsf{m}_2{\uparrow}\dfrac{\exists x(A \wedge B)}{\exists x A \wedge \exists x B} \end{array}}$$

- KSgq, a general, cut-free first-order system;

$$\mathsf{KSgq} = \mathsf{KSg} + \boxed{\begin{array}{cc} \mathsf{n}{\downarrow}\dfrac{[t \Rightarrow x]A}{\exists x A} & \mathsf{u}{\downarrow}\dfrac{\forall x(A \vee B)}{\forall x A \vee \exists x B} \end{array}}$$

- KSq an atomic cut-free first-order system.

$$\mathsf{KSq} = \mathsf{KS} + \boxed{\begin{array}{cccc} \mathsf{n}{\downarrow}\dfrac{[t \Rightarrow x]A}{\exists x A} & \mathsf{u}{\downarrow}\dfrac{\forall x(A \vee B)}{\forall x A \vee \exists x B} & \mathsf{m}_1{\downarrow}\dfrac{\exists x A \vee \exists x B}{\exists x(A \vee B)} & \mathsf{m}_2{\downarrow}\dfrac{\forall x A \vee \forall x B}{\forall x(A \vee B)} \end{array}}$$

For the rules $\mathsf{n}{\downarrow}$ and $\mathsf{n}{\uparrow}$, we insist that $x \notin FV(t)$.

All the above systems extend the equality relation from CPL with the following equations:

$$\boxed{\begin{array}{lll} \forall x A = \forall z([z \Rightarrow x]A) & \forall x \forall y A = \forall y \forall x A & \forall x B = B \\ \exists z A = \exists z([z \Rightarrow x]A) & \exists x \exists y A = \exists y \exists x A & \exists x B = B \end{array}}$$

where $x \notin FV(B)$

Definition 3.9. We say that a FOL derivation is *regular* if no variable is used in two different quantifiers.

Convention 3.10. While we often insist for lemmas and theorems that derivations be regular, unless variable hygiene is necessary for a particular proof we will often violate regularity when depicting proofs. For example, technically a contraction should have three sets of quantifiers, e.g. $\mathsf{c}\!\downarrow\frac{\exists xPx \vee \exists yPy}{\exists zPz}$, but, in practice, we would usually write this as $\mathsf{c}\!\downarrow\frac{\exists xPx \vee \exists xPx}{\exists xPx}$.

3.1.3 The Rules of Passage

We now introduce eight inference rules, the *rules of passage* (also known as *retract rules* [Brü06a]), originally described by Herbrand [Her30; Her71]. Since they are essentially used in Herbrand's work as rewriting rules, they can be considered deep inference rules *avant la lettre*.

Definition 3.11. The following eight rules are the rules of passage

$$\mathsf{r1}\!\downarrow\frac{\forall x(A \vee B)}{\forall xA \vee B} \quad \mathsf{r2}\!\downarrow\frac{\forall x(A \wedge B)}{\forall xA \wedge B} \quad \mathsf{r3}\!\downarrow\frac{\exists x(A \vee B)}{\exists xA \vee B} \quad \mathsf{r4}\!\downarrow\frac{\exists x(A \wedge B)}{\exists xA \wedge B}$$

$$\mathsf{r1}\!\uparrow\frac{\exists xA \wedge B}{\exists x(A \wedge B)} \quad \mathsf{r2}\!\uparrow\frac{\exists xA \vee B}{\exists x(A \vee B)} \quad \mathsf{r3}\!\uparrow\frac{\forall xA \wedge B}{\forall x(A \wedge B)} \quad \mathsf{r4}\!\uparrow\frac{\forall xA \vee B}{\forall x(A \vee B)}$$

where $x \in FV(B)$.

We refer to the down rules collectively as $\mathsf{RP}_\downarrow = \{\mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow, \mathsf{r3}\!\downarrow, \mathsf{r4}\!\downarrow\}$, the up rules as $\mathsf{RP}_\uparrow = \{\mathsf{r1}\!\uparrow, \mathsf{r2}\!\uparrow, \mathsf{r3}\!\uparrow, \mathsf{r4}\!\uparrow\}$ and all the rules of passage as $\mathsf{RP} = \mathsf{RP}_\downarrow \cup \mathsf{RP}_\uparrow$.

The rules of passage allow for prenexification and deprenexification of formulae. Since there are formulae whose cut-free sequent calculus proofs are non-elementarily shorter than their prenexified forms [AB16; BL94; Sta79], adding the four rules $\{\mathsf{r1}\!\uparrow, \mathsf{r2}\!\uparrow, \mathsf{r3}\!\uparrow, \mathsf{r4}\!\uparrow\}$ to a cut-free first-order system leads to significantly shorter proofs.

Proposition 3.12. *There is no elementary function bounding the length of the shortest* KSq *proof of a formula in terms of its shortest* $\mathsf{KSq} \cup \mathsf{RP}_\uparrow$ *proof.*

Proof. Immediate from the same reasoning about the size of Herbrand disjunctions in [AB16]. □

Although the 'up' rules of passage drastically change the power of proof systems from a complexity point of view, the 'down' rules of passage are all simply derivable for the cut-free FOL system KSq, with the exception of $\mathsf{r2}\!\downarrow$, which is shown to be admissible without too much difficulty.

The $\mathsf{r1}\!\downarrow$ rule is roughly equivalent to the $\mathsf{u}\!\downarrow$ rule.

Proposition 3.13. $\mathsf{r1}\!\downarrow$ *is derivable for* $\{\mathsf{u}\!\downarrow\}$, *and* $\mathsf{u}\!\downarrow$ *is derivable for* $\{\mathsf{r1}\!\downarrow, \mathsf{n}\!\downarrow\}$.

Proof.

$$\mathsf{r1}{\downarrow}\frac{\forall x(A \vee B)}{\forall xA \vee B} \quad\longrightarrow\quad \mathsf{u}{\downarrow}\frac{\forall x(A \vee B)}{\forall xA \vee {=}\dfrac{\exists xB}{B}} \qquad \mathsf{u}{\downarrow}\frac{\forall x(A \vee B)}{\forall xA \vee \exists xB} \quad\longrightarrow\quad \mathsf{r1}{\downarrow}\frac{\forall x\boxed{A \vee \mathsf{n}{\downarrow}\dfrac{B}{\exists xB}}}{\forall xA \vee \exists xB}$$

□

The r2↓ rule—an invertible rule, it should be noted—is in fact immediately derivable from a first-order medial-like up rule.

Proposition 3.14. r1↓ *is derivable for* $\{\mathsf{m}_2{\uparrow}\}$.

Proof.

$$\mathsf{r2}{\downarrow}\frac{\forall x(A \wedge B)}{\forall xA \wedge B} \quad\longrightarrow\quad \mathsf{m}_2{\uparrow}\frac{\forall x(A \wedge B)}{\forall xA \vee {=}\dfrac{\forall xB}{B}}$$

□

However, it is only admissible, not derivable for KSq.

Proposition 3.15. r1↓ *is admissible for* KSq.

Proof. If we have a proof $\overset{\Vert\,\mathsf{KSq}}{K\left\{\phi\,\dfrac{\forall x(A \wedge B)}{\forall xA \wedge B)}\right\}}$, by extending Proposition 1.52 to KSq (which is straightforward), we create:

$$\phi_l\overset{\Vert\,\mathsf{KSq}}{K\{\forall xA\}} \text{ and } \phi_r\overset{\Vert\,\mathsf{KSq}}{K\{\forall xB\}}$$

Clearly, since $x \notin FV(B)$, we can construct $\phi'_r\overset{\Vert\,\mathsf{KSq}}{K\{B\}}$ by deleting all instances of $\forall x$. We then can construct:

$$\begin{array}{c}\boxed{\phi_l\overset{\Vert\,\mathsf{KSq}}{K\{\forall xA\}} \wedge \phi'_r\overset{\Vert\,\mathsf{KSq}}{K\{B\}}} \\ \Vert\,\{\mathsf{s},\mathsf{ac}{\downarrow},\mathsf{m},\mathsf{m}_1{\downarrow},\mathsf{m}_2{\downarrow},\mathsf{aw}{\uparrow}\} \\ K\{\forall xA \wedge B\}\end{array}$$

using a slight adaptation of the second part of Proposition 1.52. Then we eliminate aw↑ using W, which works just as well in a first-order setting. □

Proposition 3.16. r3↓ *and* r4↓ *are derivable for* {n↓}, *and, dually,* r3↑ *and* r4↑ *for* {n↑}.

SKS

$$\mathsf{ai}\!\uparrow\frac{a \wedge \bar{a}}{\mathsf{f}} \qquad \mathsf{ac}\!\uparrow\frac{a}{a \wedge a} \qquad \mathsf{aw}\!\uparrow\frac{a}{\mathsf{t}}$$

SKSq

$$\mathsf{u}\!\uparrow\frac{\exists x A \wedge \forall x B}{\exists x (A \wedge B)} \qquad \mathsf{m_2}\!\uparrow\frac{\exists x (A \wedge B)}{\exists x A \wedge \exists x B}$$

$$\mathsf{n}\!\uparrow\frac{\forall x A}{[t \Rightarrow x]A} \qquad \mathsf{m_1}\!\uparrow\frac{\forall x (A \wedge B)}{\forall x A \wedge \forall x B}$$

KS

$$\mathsf{ai}\!\downarrow\frac{\mathsf{t}}{a \vee \bar{a}} \qquad \mathsf{ac}\!\downarrow\frac{a \vee a}{a} \qquad \mathsf{aw}\!\downarrow\frac{\mathsf{f}}{a}$$

$$\mathsf{s}\frac{A \wedge (B \vee C)}{(A \wedge B) \vee C} \qquad \mathsf{m}\frac{(A \wedge B) \vee (C \wedge D)}{(A \vee C) \wedge (B \vee D)}$$

KSq

$$\mathsf{n}\!\downarrow\frac{[t \Rightarrow x]A}{\exists x A} \qquad \mathsf{m_1}\!\downarrow\frac{\exists x A \vee \exists x B}{\exists x [A \vee B]}$$

$$\mathsf{u}\!\downarrow\frac{\forall x [A \vee B]}{\forall x A \vee \exists x B} \qquad \mathsf{m_2}\!\downarrow\frac{\forall x A \vee \forall x B}{\forall x [A \vee B]}$$

SKSg

$$\mathsf{i}\!\uparrow\frac{A \wedge \bar{A}}{\mathsf{f}} \qquad \mathsf{c}\!\uparrow\frac{A}{A \wedge A} \qquad \mathsf{w}\!\uparrow\frac{A}{\mathsf{t}}$$

SKSgq

$$\mathsf{u}\!\uparrow\frac{\exists x A \wedge \forall x B}{\exists x (A \wedge B)} \qquad \mathsf{n}\!\uparrow\frac{\forall x A}{[t \Rightarrow x]A}$$

KSg

$$\mathsf{i}\!\downarrow\frac{\mathsf{t}}{A \vee \bar{A}} \qquad \mathsf{c}\!\downarrow\frac{A \vee A}{A} \qquad \mathsf{w}\!\downarrow\frac{\mathsf{f}}{A}$$

$$\mathsf{s}\frac{A \wedge (B \vee C)}{(A \wedge B) \vee C}$$

KSgq

$$\mathsf{n}\!\downarrow\frac{[t \Rightarrow x]A}{\exists x A} \qquad \mathsf{u}\!\downarrow\frac{\forall x [A \vee B]}{\forall x A \vee \exists x B}$$

Figure 3.1: The "cube" of core classical proof systems. The three dimensions are cut-free/cut-full (S), atomic/general structural rules (g), and propositional/first-order (q).

Proof.

$$\mathsf{r3}\!\downarrow\frac{\exists x (A \vee B)}{\exists x A \vee B} \longrightarrow \;=\frac{\exists x \boxed{\mathsf{n}\!\downarrow\frac{A}{\exists x A} \vee B}}{\exists x A \vee B} \qquad \mathsf{r4}\!\downarrow\frac{\exists x (A \wedge B)}{\exists x A \wedge B} \longrightarrow \;=\frac{\exists x \boxed{\mathsf{n}\!\downarrow\frac{A}{\exists x A} \wedge B}}{\exists x A \wedge B}$$

□

3.1.4 Preliminary Results

We now restate and reprove certain lemmas and propositions that can almost all be found in [Brü03a] and [Brü06a]. There is no problem with their original exposition, but they will be useful to have on hand in the open deduction formalism rather than the calculus of structures.

Strong Equivalence of SKSgq and SKSq

As we showed in Proposition 1.42, the atomic CPL system SKS is strongly equivalent to the general CPL system SKSg. We can easily obtain the same strong equivalence result for the atomic FOL system SKSq and the general FOL system SKSgq, by extending Lemma 1.40 for each identity, contraction and weakening. We then need to show that $\mathsf{m}_1\!\downarrow$ and $\mathsf{m}_2\!\downarrow$ are derivable for $\{\mathsf{w}\!\downarrow, \mathsf{c}\!\downarrow\}$.

Lemma 3.17. $\mathsf{i}\!\downarrow$ *is derivable for* $\{\mathsf{ai}\!\downarrow, \mathsf{s}, \mathsf{u}\!\downarrow\}$.

Proof. We have one further inductive case than in Lemma 1.40 to consider:

$$\mathsf{i}\!\downarrow\frac{\mathsf{t}}{\forall xA \vee \exists x\bar{A}} \quad\longrightarrow\quad \mathsf{u}\!\downarrow\frac{\forall x\boxed{\mathsf{i}\!\downarrow\dfrac{\mathsf{t}}{A\vee\bar{A}}}}{\forall xA\vee\exists x\bar{A}}$$

□

Lemma 3.18. $\mathsf{c}\!\downarrow$ *is derivable for* $\{\mathsf{ac}\!\downarrow, \mathsf{m}, \mathsf{m}_1\!\downarrow, \mathsf{m}_2\!\downarrow\}$.

Proof. We have two further inductive cases than in Lemma 1.40 to consider:

$$\mathsf{c}\!\downarrow\frac{\exists xA\vee\exists xA}{\exists xA} \quad\longrightarrow\quad \mathsf{m}_1\!\downarrow\frac{\exists xA\vee\exists xA}{\exists x\boxed{\mathsf{c}\!\downarrow\dfrac{A\vee A}{A}}} \qquad \mathsf{c}\!\downarrow\frac{\forall xA\vee\forall xA}{\forall xA} \quad\longrightarrow\quad \mathsf{m}_2\!\downarrow\frac{\forall xA\vee\forall xA}{\forall x\boxed{\mathsf{c}\!\downarrow\dfrac{A\vee A}{A}}}$$

□

Lemma 3.19. $\mathsf{w}\!\downarrow$ *is derivable for* $\{\mathsf{aw}\!\downarrow\}$.

Proof. We have two further inductive cases than in Lemma 1.40 to consider:

$$\mathsf{w}\!\downarrow\frac{\mathsf{f}}{\exists xA} \quad\longrightarrow\quad =\frac{\mathsf{f}}{\exists x\boxed{\mathsf{w}\!\downarrow\dfrac{\mathsf{f}}{A}}} \qquad \mathsf{w}\!\downarrow\frac{\mathsf{f}}{\forall xA} \quad\longrightarrow\quad =\frac{\mathsf{f}}{\forall x\boxed{\mathsf{w}\!\downarrow\dfrac{\mathsf{f}}{A}}}$$

□

Lemma 3.20. $\mathsf{m}_1\!\downarrow$ *and* $\mathsf{m}_2\!\downarrow$ *are derivable for* $\{\mathsf{w}\!\downarrow, \mathsf{c}\!\downarrow\}$.

Proof.

$$\mathsf{m}_1\!\downarrow\frac{\exists xA\vee\exists xB}{\exists x(A\vee B)} \quad\longrightarrow\quad \mathsf{c}\!\downarrow\frac{\exists x\boxed{=\dfrac{A}{A\vee\mathsf{w}\!\downarrow\dfrac{\mathsf{f}}{B}}}\vee\exists x\boxed{=\dfrac{B}{B\vee\mathsf{w}\!\downarrow\dfrac{\mathsf{f}}{A}}}}{\exists x(A\vee B)}$$

$$\mathsf{m_2}\!\downarrow \frac{\forall x A \vee \forall x B}{\forall x (A \vee B)} \quad \longrightarrow \quad \mathsf{c}\!\downarrow \frac{\forall x \boxed{= \frac{A}{A \vee \mathsf{w}\!\downarrow \frac{\mathsf{f}}{B}}} \vee \forall x \boxed{= \frac{B}{B \vee \mathsf{w}\!\downarrow \frac{\mathsf{f}}{A}}}}{\forall x (A \vee B)}$$

□

Proposition 3.21. *The atomic system* SKSq *and the general* SKSgq *are strongly equivalent.*

Proof. Follows from Propositions 3.17, 3.18, 3.20 and 3.19, as well as their dualised versions. □

Admissibility of n↑

An important property of the cut-free system KSgq is that n↑, universal quantifier instantiation, is directly admissible, without having to resort to admissibility via cut elimination, and in a way that does not have any significant proof complexity cost.

Proposition 3.22. n↑ *is admissible for* KSgq

Proof. We show that we can n↑ permute up past all other rules. The five reductions we need to consider are the following:

$$\rho - \mathsf{n}\!\uparrow_1 : \quad = \frac{K\{\forall x A\}\left\{\rho \frac{B}{C}\right\}}{K\left\{\mathsf{n}\!\uparrow \frac{\forall x A}{[t \Rightarrow x]A}\right\}\{C\}} \quad \longrightarrow \quad = \frac{K\left\{\mathsf{n}\!\uparrow \frac{\forall x A}{[t \Rightarrow x]A}\right\}\{B\}}{K\{[t \Rightarrow x]A\}\left\{\rho \frac{B}{C}\right\}}$$

$$\rho - \mathsf{n}\!\uparrow_2 : \quad \mathsf{n}\!\uparrow \frac{\forall x K\left\{\rho \frac{A}{B}\right\}}{[t \Rightarrow x]K\{B\}} \quad \longrightarrow \quad \mathsf{n}\!\uparrow \frac{\forall x K\{A\}}{[t \Rightarrow x]K\left\{\rho' \frac{A}{B}\right\}}$$

$$\rho - \mathsf{n}\!\uparrow_3 : \quad \rho \frac{K\{\forall x A\}}{K'\left\{\mathsf{n}\!\uparrow \frac{\forall x A}{[t \Rightarrow x]A}\right\}} \quad \longrightarrow \quad \rho \frac{K\left\{\mathsf{n}\!\uparrow \frac{\forall x A}{[t \Rightarrow x]A}\right\}}{K'\{[t \Rightarrow x]A\}}$$

$$\mathsf{c}\!\downarrow - \mathsf{n}\!\uparrow : \quad \mathsf{c}\!\downarrow \frac{K\{\forall x A\} \vee K\{\forall x A\}}{K\left\{\mathsf{n}\!\uparrow \frac{\forall x A}{[t \Rightarrow x]A}\right\}} \quad \longrightarrow \quad \mathsf{c}\!\downarrow \frac{K\left\{\mathsf{n}\!\uparrow \frac{\forall x A}{[t \Rightarrow x]A}\right\} \vee K\left\{\mathsf{n}\!\uparrow \frac{\forall x A}{[t \Rightarrow x]A}\right\}}{K\{[t \Rightarrow x]A\}}$$

$$\mathsf{u}\!\downarrow - \mathsf{n}\!\uparrow : \quad \mathsf{u}\!\downarrow \frac{\forall x (A \vee B)}{\mathsf{n}\!\uparrow \frac{\forall x A}{[t \Rightarrow x]A} \vee \exists x B} \quad \longrightarrow \quad \mathsf{n}\!\uparrow \frac{\forall x (A \vee B)}{[t \Rightarrow x]A \vee \mathsf{n}\!\downarrow \frac{[t \Rightarrow x]B}{\exists x B}}$$

Clearly, the $\rho - \mathsf{n}\!\uparrow_1$ reduction is always straightforward. $\rho - \mathsf{n}\!\uparrow_2$ is straightforward in most cases, except if $\rho = \mathsf{n}\!\downarrow$, where we should be careful with substitutions:

$$\mathsf{n}\!\uparrow \frac{\forall x K \left\{ \mathsf{n}\!\downarrow \dfrac{[t_2 \Rightarrow y]A}{\exists y A} \right\}}{[t_1 \Rightarrow x] K \{\exists y A\}} \quad \longrightarrow \quad = \frac{= \dfrac{\mathsf{n}\!\uparrow \dfrac{\forall x K\{[t_2 \Rightarrow y]A\}}{[t_1 \Rightarrow x]K\{[t_2 \Rightarrow y]A\}}}{[t_1 \Rightarrow x]K \left\{ \mathsf{n}\!\downarrow \dfrac{[[t_1 \Rightarrow x]t_2 \Rightarrow y][t_1 \Rightarrow x]A}{\exists y [t_1 \Rightarrow x]A} \right\}}}{[t_1 \Rightarrow x](K\{\exists y A\})}$$

$\rho - \mathsf{n}\!\uparrow_3$ is also straightforward except if $\rho = \mathsf{n}\!\downarrow$, where we should be dually careful with substitutions:

$$\mathsf{n}\!\downarrow \frac{[t_1 \Rightarrow x]K\{\forall y A\}}{\exists x K \left\{ \mathsf{n}\!\uparrow \dfrac{\forall y A}{[t_2 \Rightarrow y]A} \right\}} \quad \longrightarrow \quad = \frac{[t_1 \Rightarrow x]K\{\forall y A\}}{= \dfrac{[t_1 \Rightarrow x]K \left\{ \mathsf{n}\!\uparrow \dfrac{\forall y [t_1 \Rightarrow x]A}{[[t_1 \Rightarrow x]t_2 \Rightarrow y][t_1 \Rightarrow x]A} \right\}}{\mathsf{n}\!\downarrow \dfrac{[t_1 \Rightarrow x]K\{[t_2 \Rightarrow y]A\}}{\exists x K\{[t_2 \Rightarrow y]A\}}}}$$

□

Elimination of the "up" rules as cut elimination

In propositional logic, cut elimination is performed by reducing cut to atomic form with switches and then eliminating the atomic cuts. Unfortunately, this strategy does not cleanly lift to first-order logic, since, to reduce cut to atomic form, other "up" rules are needed. Instead, we show that we can reduce the instances of cut in a proof to two types: *closed* atomic cuts and *closed quantifier cuts*. Quantifier cuts are those where the principal formula has a quantifier as its outermost connective; closed cuts are those with no free variables.

Definition 3.23. We say that an instance of a rule ρ in a FOL proof is closed if its premise and conclusion are both closed formulae.

Proposition 3.24. *Let* $\begin{array}{c}\phi \,\|\, \mathsf{SKSq} \\ A\end{array}$. *Then we can construct* $\begin{array}{c}\phi' \,\|\, \mathsf{KSq} \cup \{\mathsf{i}\!\uparrow\} \\ A\end{array}$, *where every instance of* $\mathsf{i}\!\uparrow$ *is closed.*

$$\begin{array}{c}\phi \,\|\, \mathsf{SKSq} \\ A\end{array} \longrightarrow \begin{array}{c}\phi' \,\|\, \mathsf{KSq} \cup \{\mathsf{i}\!\uparrow\} \\ A\end{array}$$

Proof. Similar to Proposition 1.46, except we need to make sure that the $\mathsf{i}\!\uparrow$ is closed. We do so by including a context around $\rho\downarrow$ s.t. $K\{A\}$ is a closed formula.

$$K\left\{\rho\uparrow\frac{A}{B}\right\} \quad\longrightarrow\quad =\dfrac{K\{A\}}{\mathsf{s}\dfrac{K\{A\}\wedge \mathsf{i}\!\downarrow\dfrac{\mathsf{t}}{\bar{K}\left\{\rho\!\downarrow\dfrac{\bar{B}}{\bar{A}}\right\}\vee K\{B\}}}{\mathsf{i}\!\uparrow\dfrac{K\{A\}\wedge\bar{K}\left\{\bar{A}\right\}}{\mathsf{f}}\vee K\{B\}}}$$

□

Definition 3.25. We define the rules qi↓ and qi↑ to be instances of i↓ and i↑ where the principal formulae have a quantifier as their outermost connective:

$$\mathsf{qi}\!\downarrow\frac{\mathsf{t}}{\exists xA\vee\forall x\bar{A}} \quad\text{and}\quad \mathsf{qi}\!\uparrow\frac{\exists xA\wedge\forall x\bar{A}}{\mathsf{f}}$$

Lemma 3.26.

$$\begin{array}{c}\phi\,\|\,\mathsf{KSq}\cup\{\mathsf{i}\uparrow\}\\ A\end{array} \longrightarrow \begin{array}{c}\phi'\,\|\,\mathsf{KSq}\cup\{\mathsf{ai}\uparrow,\mathsf{qi}\uparrow\}\\ A\end{array}$$

with all ai↑, qi↑ *closed.*

Proof. Clearly, i↑ is derivable for {ai↑, qi↑, s}. □

We can now go one step further and separate proofs into two parts: a lower part containing only quantifier cuts, and a top part containing all the other inference rules, including atomic cuts. This makes cut elimination a lot more manageable, and also will allow for useful translations later on in the thesis.

Lemma 3.27. *For any first-order formula context $K\{\ \}$ and any formula A, with no free variables in A bound by $K\{\ \}$, there are derivations*

$$\begin{array}{c}K\{\mathsf{t}\}\wedge A\\ \|\,\{\mathsf{s},\mathsf{n}\uparrow,\mathsf{u}\uparrow\}\\ K\{A\}\end{array} \quad\textit{and}\quad \begin{array}{c}K\{A\}\\ \|\,\{\mathsf{s},\mathsf{n}\downarrow,\mathsf{u}\downarrow\}\\ K\{\mathsf{f}\}\vee A\end{array}$$

.

Proof. We have two additional inductive steps to prove, in addition to those in 1.53:

$$=\dfrac{\dfrac{K\{\mathsf{t}\}}{\forall xK'\{\mathsf{t}\}}\wedge A}{=\dfrac{\forall x'\boxed{\begin{array}{c}\mathsf{n}\!\uparrow\dfrac{\forall xK'\{\mathsf{t}\}}{[x'\Rightarrow x]K'\{\mathsf{t}\}}\wedge A\\ IH\,\|\,\{\mathsf{s},\mathsf{n}\uparrow,\mathsf{u}\uparrow\}\\ [x'\Rightarrow x]K'\{A\}\end{array}}}{K'\{A\}}} \quad\text{and}\quad \mathsf{u}\!\uparrow\dfrac{\dfrac{K\{\mathsf{t}\}}{\exists xK'\{\mathsf{t}\}}\wedge =\dfrac{A}{\forall xA}}{=\dfrac{\exists x\boxed{\begin{array}{c}K'\{\mathsf{t}\}\wedge A\\ IH\,\|\,\{\mathsf{s},\mathsf{n}\uparrow,\mathsf{u}\uparrow\}\\ K'\{A\}\end{array}}}{K\{A\}}}$$

□

Lemma 3.28. *Let* $\begin{array}{c}\phi \, \| \, \mathsf{KSq} \cup \{\mathsf{ai}\uparrow, \mathsf{qi}\uparrow\} \\ A\end{array}$ *be a proof with all* $\mathsf{ai}\uparrow, \mathsf{qi}\uparrow$ *closed. Then we can construct a proof*

$$\begin{array}{c} \phi' \, \Big\| \, \mathsf{KSq} \cup \{\mathsf{ai}\uparrow\} \\ = \dfrac{A \vee \boxed{\mathsf{qi}\uparrow \dfrac{\exists x A_1 \wedge \forall x \bar{A}_1}{\mathsf{f}}} \vee \ldots \vee \boxed{\mathsf{qi}\uparrow \dfrac{\exists x A_n \wedge \forall x \bar{A}_n}{\mathsf{f}}}}{A} \end{array}$$

where all the $\mathsf{ai}\uparrow, \mathsf{qi}\uparrow$ *are still closed.*

Proof. This is essentially a one-way version of Lemma 1.54, for first-order logic, using Lemma 3.27 instead of Lemma 1.53. ☐

These lemmas lead us to the following proposition, which sets us up for the next section.

Proposition 3.29. *Every proof* SKSq *proof* ϕ *of* A *can be separated into a quantifier-cut-free top half, with parallel closed quantifier-cuts in the bottom half:*

$$\begin{array}{c} \phi \, \| \, \mathsf{SKSq} \\ A \end{array} \longrightarrow \begin{array}{c} \| \, \mathsf{KSq} \cup \{\mathsf{ai}\uparrow\} \\ A' \\ \| \, \{\mathsf{qi}\uparrow\} \\ A \end{array}$$

Proof. By Lemmas 3.24, 3.26 and 3.28. ☐

3.2 First-Order Merge contractions

We now introduce first-order systems with merge contractions. Unlike for propositional logic, we do not have a particular use for them in the course of normalization for first-order logic. However, we think it is worth extending the definitions and basic theory to first-order logic, as we expect further research in this area.

3.2.1 SKSq_4 and First-order Contractive Derivations

Before defining merge contractions for first-order logic, we need to extend the propositional system with mostly "medial-shaped" rules to FOL. Since the first-order rules in SKSq are in fact well suited for merges, we simply extend SKS_4 with them.

Definition 3.30. We define the variant of SKSq with medial-shaped propositional rules SKSq_4 as follows.

$$\mathsf{SKSq}_4 = \mathsf{SKS}_4 + \begin{array}{cccc} \mathsf{n}{\downarrow}\dfrac{[t \Rightarrow x]A}{\exists x A} & \mathsf{u}{\downarrow}\dfrac{\forall x(A \vee B)}{\forall x A \vee \exists x B} & \mathsf{m}_1{\downarrow}\dfrac{\exists x A \vee \exists x B}{\exists x(A \vee B)} & \mathsf{m}_2{\downarrow}\dfrac{\forall x A \vee \forall x B}{\forall x(A \vee B)} \\[2ex] \mathsf{n}{\uparrow}\dfrac{\forall x A}{[t \Rightarrow x]A} & \mathsf{u}{\uparrow}\dfrac{\exists x A \wedge \forall x B}{\exists x(A \wedge B)} & \mathsf{m}_1{\uparrow}\dfrac{\forall x(A \wedge B)}{\forall x A \wedge \forall x B} & \mathsf{m}_2{\uparrow}\dfrac{\exists x(A \wedge B)}{\exists x A \wedge \exists x B} \end{array}$$

In addition, we also restrict the equality relation as we do for SKS_4

Definition 3.31. The set of *first-order contractive derivations*, usually just referred to as contractive derivations if unambiguous, is the minimal set of $\mathsf{C1}_\downarrow = \{\mathsf{ac}{\downarrow}, \mathsf{m}, \vee\mathsf{c}{\downarrow}, \mathsf{m}_1{\downarrow}, \mathsf{m}_2{\downarrow}\}$ derivations that satisfy each of the three properties below:

CD1 A formula $A \vee B$ is a contractive derivation.

CD2 $=\dfrac{\mathsf{f} \vee \mathsf{f}}{\mathsf{f}}, =\dfrac{\mathsf{t} \vee \mathsf{t}}{\mathsf{t}}$ and $\mathsf{ac}{\downarrow}\dfrac{a \vee a}{a}$ are contractive derivations.

CD3 $=\dfrac{\mathsf{f} \vee A}{A}$ and $=\dfrac{A \vee \mathsf{f}}{A}$ are contractive derivations.

CD4 If $\begin{array}{c} A_1 \vee B_1 \\ \phi_1 \| \mathsf{C1}_\downarrow \\ C_1 \end{array}$ and $\begin{array}{c} A_2 \vee B_2 \\ \phi_2 \| \mathsf{C1}_\downarrow \\ C_2 \end{array}$ are contractive derivations then both

$$\mathsf{m}\dfrac{(A_1 \wedge A_2) \vee (B_1 \wedge B_2)}{\begin{array}{c} A_1 \vee B_1 \\ \phi_1 \| \mathsf{C1}_\downarrow \\ C_1 \end{array} \wedge \begin{array}{c} A_2 \vee B_2 \\ \phi_2 \| \mathsf{C1}_\downarrow \\ C_2 \end{array}} \quad \text{and} \quad \vee\mathsf{c}{\downarrow}\dfrac{(A_1 \vee A_2) \vee (B_1 \vee B_2)}{\begin{array}{c} A_1 \vee B_1 \\ \phi_1 \| \mathsf{C1}_\downarrow \\ C_1 \end{array} \vee \begin{array}{c} A_2 \vee B_2 \\ \phi_2 \| \mathsf{C1}_\downarrow \\ C_2 \end{array}}$$

are contractive derivations.

CD5 If $\begin{array}{c} A \vee B \\ \phi \| \mathsf{C1}_\downarrow \\ C \end{array}$ is a contractive derivation, then both

$$\mathsf{m}_1{\downarrow}\dfrac{\forall x A \vee \forall x B}{\forall x \begin{array}{c} A \vee B \\ \phi \| \mathsf{C1}_\downarrow \\ C \end{array}} \quad \text{and} \quad \mathsf{m}_2{\downarrow}\dfrac{\exists x A \vee \exists x B}{\exists x \begin{array}{c} A \vee B \\ \phi \| \mathsf{C1}_\downarrow \\ C \end{array}}$$

are contractive derivations.

The set of *first-order cocontractive derivations* is defined dually.

We also define first-order merge sets:

Definition 3.32. Given two first-order formulae A, B and $\star \in \{\vee, \wedge\}$, we define their *$\star$-merge set* $M_\star(A, B)$ as the minimum set that satisfies the following conditions

M1 For any A and B, $A \star B \in M_\star(A,B)$.

M2 For any atom or unit a, $a \in M_\star(a,a)$.

M3 For any A, $A \in M_\vee(A,\mathsf{f}), A \in M_\vee(\mathsf{f},A), A \in M_\wedge(A,\mathsf{t})$ and $A \in M_\wedge(\mathsf{t},A)$.

M4 For $\alpha \in \{\vee,\wedge\}$, if $C_1 \in M_\star(A_1,B_1)$ and $C_2 \in M_\star(A_2,B_2)$, then $C_1\ \alpha\ C_2 \in M_\star(A_1\ \alpha\ A_2, B_1\ \alpha\ B_2)$.

M5 For $Q \in \{\forall,\exists\}$, if $C \in M_\star(A,B)$ then $QxC \in M_\star(QxA,QxB)$.

Remark 3.33. Note that if A,B,C are quantifier-free, then whether $C \in M_\vee(A,B)$ or not cannot depend on **M5**. We therefore do not alter our notation for merges in first-order logic. On the other hand, not all first-order merges need use **M5**. However, if $C \in M_\vee(A,B)$ without any use of **M5**, then C is in $M_\vee(A,B)$ for "propositional" reasons.

Proposition 3.34. *There exists a first-order contractive derivation* $\begin{array}{c} A \vee B \\ \phi \Vert \mathsf{c1}_\downarrow \\ C \end{array}$ *iff* $C \in M_\vee(A,B)$. *Dually there exists a cocontractive derivation* $\begin{array}{c} C \\ \phi \Vert \mathsf{c1}_\uparrow \\ A \wedge B \end{array}$ *iff* $C \in M_\wedge(A,B)$.

Proof. Simple extension of Proposition 2.30 □

We can now define a first-order proof system with merge (co)contractions. Again, since the instances of propositional merge (co)contraction are exactly the quantifier-free instances of first-order merge (co)contraction, we use the same notation as in propositional logic.

Definition 3.35. We define the first-order system with merge contractions:

$$\mathsf{SKSmq} = \mathsf{SKSm} + \boxed{\begin{array}{cc} \mathsf{n}\!\downarrow \dfrac{[t \Rightarrow x]A}{\exists x A} & \mathsf{u}\!\downarrow \dfrac{\forall x(A \vee B)}{\forall x A \vee \exists x B} \\[2ex] \mathsf{n}\!\uparrow \dfrac{\forall x A}{[t \Rightarrow x]A} & \mathsf{u}\!\uparrow \dfrac{\exists x A \wedge \forall x B}{\exists x(A \wedge B)} \end{array}}$$

We extend the maps from the systems with merges to the regular systems in the obvious way: $\mu : \mathsf{SKSm} \to \mathsf{SKS}$ becomes $\mu : \mathsf{SKSmq} \to \mathsf{SKSmq}$, the maps $\nu_{\max}, \nu_{\min} : \mathsf{SKS} \to \mathsf{SKSm}$ become $\nu_{\max}, \nu_{\min} : \mathsf{SKSq} \to \mathsf{SKSmq}$

3.2.2 Reductions for first-order merges

We now show that the lemmas that allow us to permute merge contractions past other inference rules extend to first-order logic. We repeat the statements of the lemmas, even though they are the same as the propositional case, if interpreted as referring to first-order formulas and merges. The proofs usually involve just one extra case: we will only give this, not repeating the propositional parts.

SKS_4

$$\mathsf{s}\!\uparrow\frac{(A\wedge B)\wedge(C\vee D)}{(A\wedge C)\vee(B\wedge D)}\qquad \wedge\mathsf{c}\!\uparrow\frac{(A\wedge B)\wedge(C\wedge D)}{(A\wedge C)\wedge(B\wedge D)}$$

$$\mathsf{ai}\!\uparrow\frac{a\wedge\bar{a}}{\mathsf{f}}\qquad \mathsf{ac}\!\uparrow\frac{a}{a\wedge a}\qquad \mathsf{aw}\!\uparrow\frac{a}{\mathsf{t}}$$

SKSq_4

$$\mathsf{u}\!\uparrow\frac{\exists xA\wedge\forall xB}{\exists x(A\wedge B)}\qquad \mathsf{m}_2\!\uparrow\frac{\exists x(A\wedge B)}{\exists xA\wedge\exists xB}$$

$$\mathsf{n}\!\uparrow\frac{\forall xA}{[t\Rightarrow x]A}\qquad \mathsf{m}_1\!\uparrow\frac{\forall x(A\wedge B)}{\forall xA\wedge\forall xB}$$

KS_4

$$\mathsf{ai}\!\downarrow\frac{\mathsf{t}}{a\vee\bar{a}}\qquad \mathsf{ac}\!\downarrow\frac{a\vee a}{a}\qquad \mathsf{aw}\!\downarrow\frac{\mathsf{f}}{a}$$

$$\mathsf{s}\!\downarrow\frac{(A\vee B)\wedge(C\vee D)}{(A\wedge C)\vee(B\vee D)}\qquad \vee\mathsf{c}\!\downarrow\frac{(A\vee B)\vee(C\vee D)}{(A\vee C)\vee(B\vee D)}$$

$$\mathsf{m}\frac{(A\wedge B)\vee(C\wedge D)}{(A\vee C)\wedge(B\vee D)}$$

KSq_4

$$\mathsf{n}\!\downarrow\frac{[t\Rightarrow x]A}{\exists xA}\qquad \mathsf{m}_1\!\downarrow\frac{\exists xA\vee\exists xB}{\exists x(A\vee B)}$$

$$\mathsf{u}\!\downarrow\frac{\forall x(A\vee B)}{\forall xA\vee\exists xB}\qquad \mathsf{m}_2\!\downarrow\frac{\forall xA\vee\forall xB}{\forall x(A\vee B)}$$

SKSm

$$\mathsf{s}\!\uparrow\frac{(A\wedge B)\wedge(C\vee D)}{(A\wedge C)\vee(B\wedge D)}$$

$$\mathsf{ai}\!\uparrow\frac{a\wedge\bar{a}}{\mathsf{f}}\qquad \mathsf{mc}\!\uparrow\frac{C}{A\wedge B}\qquad \mathsf{aw}\!\uparrow\frac{a}{\mathsf{t}}$$

SKSmq

$$\mathsf{u}\!\uparrow\frac{\exists xA\wedge\forall xB}{\exists x(A\wedge B)}\qquad \mathsf{n}\!\uparrow\frac{\forall xA}{[t\Rightarrow x]A}$$

KSm

$$\mathsf{ai}\!\downarrow\frac{\mathsf{t}}{a\vee\bar{a}}\qquad \mathsf{mc}\!\downarrow\frac{A\vee B}{C}\qquad \mathsf{aw}\!\downarrow\frac{\mathsf{f}}{A}$$

$$\mathsf{s}\!\downarrow\frac{(A\vee B)\wedge(C\vee D)}{(A\wedge C)\vee(B\vee D)}\qquad \mathsf{m}\frac{(A\wedge B)\vee(C\wedge D)}{(A\vee C)\wedge(B\vee D)}$$

KSmq

$$\mathsf{n}\!\downarrow\frac{[t\Rightarrow x]A}{\exists xA}\qquad \mathsf{u}\!\downarrow\frac{\forall x(A\vee B)}{\forall xA\vee\exists xB}$$

Figure 3.2: The "cube" of medial-like/merge proof systems. The three dimensions are cut-free/cut-full (S), medial-shape/merge contractions (4/m), and propositional/first-order (q). Taken together with the proof systems of Figure 3.1, a hypercube of proof systems is formed.

Definition 3.36. Let $K\{D\} \in M_\vee(A,B)$. We say that D is a K-subformula in A or B in the following cases:

M1 If $K\{D\} \equiv K'\{D\} \vee B$ and $A \equiv K'\{D\}$ then D is a K-subformula of A. If $K\{D\} \equiv A \vee K'\{D\}$ and $B \equiv K'\{D\}$ then D is a K-subformula of B.

M3 If $K\{D\} \equiv A$, then D is a K-subformula of A

M4 If $K_1\{D_1\} \in M_\vee(A_1,B_1)$ and D_1 is a K_1-subformula of A_1 (resp. B_1), then for all $C_2 \in M_\vee(A_2,B_2)$, D_1 is a K-subformula of $A_1 \ \alpha \ A_2$ (resp. $B_1 \ \alpha \ B_2$), where $K\{\ \} = K_1\{\ \} \ \alpha \ C_2$.

M5 If $K\{D\} \in M_\vee(A,B)$ with D a K-subformula of A, then we have that D is a QxK-subformula of QxA.

Lemma 3.37. *Let A,B,C be formulae with $C \in M^P_\vee(A,B)$. Let $K_C\{\ \}$ and P_C be s.t. $C \equiv K_C\{P_C\}$ and P_C is not a K_C-subformula of A or B. Then we can find contexts $K_A\{\ \},K_B\{\ \}$ and formulae P_A,P_B s.t. A factorises as $K_A\{P_A\}$, B as $K_B\{P_B\}$, $P_C \in M_\vee(P_A,P_B)$ and for any Q_A,Q_B, if $Q_C \in M_\vee(Q_A,Q_B)$, then*

$$K_C\{Q_C\} \in M_\vee(K_A\{Q_A\},K_B\{Q_B\})\,.$$

Alternatively, we can always find $K_A\{\ \}$, $K_B\{\ \}$, P_A and P_B s.t.

$$\mathsf{mc}{\downarrow}\frac{A \vee B}{K_C\{P_C\}} \quad \equiv \quad \mathsf{mc}{\downarrow}\frac{K_A\{P_A\} \vee K_B\{P_B\}}{K_C\{P_C\}}$$

and, for any $Q_C \in M_\vee(Q_A,Q_B)$,

$$\mathsf{mc}{\downarrow}\frac{K_A\{Q_A\} \vee K_B\{Q_B\}}{K_C\{Q_C\}}$$

is a valid instance of $\mathsf{mc}{\downarrow}$.

Proof. Assume $K_C\{\ \} = QxK'_C\{\ \}$. Again, since P_C is not a K_C-subformula of A or B, we know that $A \neq \mathsf{f} \neq B$. Therefore we must have $C \in M^P_\vee(A,B)$ by virtue of **M5** with $A \equiv QxA'$, $B \equiv QxB'$ and $K'_C\{P_C\} \in M_\vee(A',B')$.

If $K'_C\{P_C\} \equiv A' \vee B'$ then, by the same reasoning as before, it must be the case that $P_C \equiv A' \vee B'$. Therefore we can factorise A as $Qx\{A'\}$ and B as $Qx\{B'\}$, with all the conditions clearly holding.

If $K'_C\{P_C\} \neq A \vee B$ then $K'_C\{P_C\} \in M^P_\vee(A,B)$ and, by the IH, we have $A' \equiv K_{A'}\{P_A\}$ and $B' \equiv K_{B'}\{P_B\}$ with all the appropriate conditions and we can factorise A as $QxK_{A'}\{P_A\}$ and B as $QxB' \equiv K_{B'}\{P_B\}$. Again, it is straightforward to see that all the conditions hold. ☐

Lemma 3.38. *Assume we have $A \in M_\vee(A_1,A_2), B \in M_\vee(B_1,B_2)$ and $C \in M^P_\vee(A,B)$. Then we can find C_1,C_2 s.t. $C_1 \in M^P_\vee(A_1,B_1)$ and $C_2 \in M^P_\vee(A_2,B_2)$ s.t. $C \in M^P_\vee(C_1,C_2)$. Alternatively, we always can find C_1 and C_2 s.t. the following rewrite is valid.*

$$\mathsf{mc}{\downarrow}\frac{(A_1 \vee B_1) \vee (A_2 \vee B_2)}{\mathsf{mc}{\downarrow}\dfrac{A \vee B}{C}} \quad \longrightarrow \quad \mathsf{mc}{\downarrow}\frac{\mathsf{mc}{\downarrow}\dfrac{A_1 \vee B_1}{C_1} \vee \mathsf{mc}{\downarrow}\dfrac{A_2 \vee B_2}{C_2}}{C}$$

Proof. If $C \equiv QxC'$, then the cases where any of A, A_1, A_2, B, B_1 or B_2 are f are as before. If not, then it must be the case that $A \equiv QxA'$ and $B \equiv QxB'$ and that $A_1 \equiv QxA_1'$, $A_2 \equiv QxA_2'$, $B_1 \equiv QxB_1'$ and $B_2 \equiv QxB_2'$. We can therefore use the IH to find C_1' and C_2', setting $C_1 \equiv QxC_1$ and $C_2 \equiv QxC_2$. □

Lemma 3.39. *Assume we have $B \in M^P_\vee(A_1, A_2)$ and $B \in M^P_\wedge(C_1, C_2)$. Then we can find B_1, B_2, B_3, B_4 s.t. $A_1 \in M^P_\wedge(B_1, B_2)$, $A_2 \in M^P_\wedge(B_3, B_4)$, $C_1 \in M^P_\vee(B_1, B_3)$ and $C_2 \in M^P_\vee(B_2, B_4)$. Alternatively, we can always find B_1, B_2, B_3, B_4 s.t. the following rewrite is valid (although in practice the medial would be absorbed into one of the merge (co)contractions):*

$$\mathsf{mc}{\uparrow}\frac{\mathsf{mc}{\downarrow}\dfrac{A_1 \vee A_2}{B}}{C_1 \wedge C_2} \quad \longrightarrow \quad \mathsf{m}\frac{\mathsf{mc}{\uparrow}\dfrac{A_1}{B_1 \wedge B_2} \vee \mathsf{mc}{\uparrow}\dfrac{A_2}{B_3 \wedge B_4}}{\mathsf{mc}{\downarrow}\dfrac{B_1 \vee B_3}{C_1} \wedge \mathsf{mc}{\downarrow}\dfrac{B_2 \vee B_4}{C_2}}$$

Proof. If $B \equiv QxB'$, then the cases where A_1 or A_2 are f and C_1 or C_2 are as before. If not, then it must be the case that $A_1 \equiv QxA_1'$, $A_2 \equiv QxA_2'$, $C_1 \equiv QxC_1'$ and $C_2 \equiv QxC_2'$. We can therefore use the IH to find B_1', B_2', B_3' and B_4', setting each B_i as QxB_i'. □

Lemma 3.40. *Let $[t \Rightarrow x]C \in M_\vee(A, B)$. Then we can find A' and B' s.t. $A \equiv [t \Rightarrow x]A'$, $B \equiv [t \Rightarrow x]B'$ and $C \in M_\vee(A', B')$.*

Proof. Two cases we can do instantly:

M1 Assume $[t \Rightarrow x]C \equiv A \vee B$. Then $C \equiv [C_1.C_2]$, and we can set $A' \equiv C_1$ and $B' \equiv C_2$.

M3 Assume, without loss of generality (WLOG), that $[t \Rightarrow x]C \equiv A$ and $B \equiv$ f. We set $A' \equiv C$, $B' \equiv$ f.

If neither of these two cases hold, we proceed by induction on C, considering each of the remaining three cases for $[t \Rightarrow x]C \in M_\vee(A, B)$:

M2 Assume $[t \Rightarrow x]C \equiv A \equiv B \equiv P(t_1, \ldots, t_n)$. Then $C \equiv P(t_1', \ldots, t_n')$, with $[t \Rightarrow x]t_i' \equiv t_i$. Thus we can set $A' \equiv B' \equiv P(t_1', \ldots, t_n')$.

M4 Assume $[t \Rightarrow x]C \equiv (C_1 \; \alpha \; C_2)$, $A \equiv (A_1 \; \alpha \; A_2)$, $B \equiv (B_1 \; \alpha \; B_2)$, $C_1 \in M_\vee(A_1, B_1)$ and $C_2 \in M_\vee(A_2, B_2)$. Clearly, $C_1 \equiv [t \Rightarrow x]C_1'$ and $C_2 \equiv [t \Rightarrow x]C_2'$, with $C \equiv (C_1' \; \alpha \; C_2')$. Thus, by the IH, we can find A_1', A_2', B_1', B_2' s.t. $C_1' \in M_\vee(A_1', B_1')$ and $C_2' \in M_\vee(A_2', B_2')$. Thus $C \equiv (C_1' \; \alpha \; C_2') \in M_\vee((A_1' \; \alpha \; A_2'), (B_1' \; \alpha \; B_2')) = M_\vee(A', B')$.

M5 Assume $[t \Rightarrow x]C \equiv QyC_1$, $A \equiv QxA_1$, $B \equiv QxB_1$ and $C_1 \in M_\vee(A_1, B_1)$. If $x = y$ then $[t \Rightarrow x]C \equiv C$ and there is nothing to do. If $x \neq y$, then $[t \Rightarrow x]C \equiv Qy[t \Rightarrow x]C_1$, and, by the IH, we can find A_1' and B_1' s.t. $C_1 \in M_\vee(A_1', B_1')$. Therefore, $C \equiv QyC_1 \in M_\vee(QyA_1', QyB_1')$.

□

We can now reprove the theorem that states that merge contractions can be permuted past all other rules, if they are in the scope of the conclusion. Again, we repeat the statement of the theorem, but only include the extra cases for the proof.

Theorem 3.41. *Assume we have a merge contraction above a rule $\rho \in \mathsf{SKSmq}$ in a context, with P_C not a K_C-subformula of A or B:*

$$\mathsf{mc}{\downarrow}\frac{A \vee B}{K_C\left\{\rho\,\dfrac{P_C}{Q_C}\right\}}$$

Then we can find sentences P_A, P_B, Q_A, Q_B, contexts $K_A\{\ \}, K_B\{\ \}$ and derivations $\phi_1, \phi_2 \in \mathsf{SKSmq}$ s.t. the following is a valid derivation:

$$\mathsf{mc}{\downarrow}\frac{K_A\left\{\begin{array}{c} P_A \\ \phi_1 \| \\ Q_A \end{array}\right\} \vee K_B\left\{\begin{array}{c} P_B \\ \phi_2 \| \\ Q_B \end{array}\right\}}{K_C\{Q_C\}}$$

- If $\rho = \mathsf{n}{\downarrow}$ then we have the general case below to consider:

$$\mathsf{mc}{\downarrow}\frac{K_A\{P_A\} \vee K_B\{P_B\}}{K_C\left\{\mathsf{n}{\downarrow}\,\dfrac{[t \Rightarrow x]P_C}{\exists x P_C}\right\}}$$

  By Lemma 3.40 we can find P_A' and P_B' s.t.

$$\mathsf{mc}{\downarrow}\frac{K_A\left\{\mathsf{n}{\downarrow}\,\dfrac{P_A}{\exists x P_A'}\right\} \vee K_B\left\{\mathsf{n}{\downarrow}\,\dfrac{P_B}{\exists x P_B'}\right\}}{K_C\{\exists x P_C\}}$$

  is valid.

- If $\rho = \mathsf{n}{\uparrow}$ then we again have a general case to consider:

$$\mathsf{mc}{\downarrow}\frac{K_A\{P_A\} \vee K_B\{P_B\}}{K_C\left\{\rho\,\dfrac{\forall x P_C}{[t \Rightarrow x]P_C}\right\}}$$

  Since $\forall x P_C$ is not a K_C-subformula of A or B, it must be the case that $P_A \equiv \forall x P_A'$ and $P_B \equiv \forall x P_B'$.

- If $\rho = \mathsf{u}{\downarrow}$ then, since P_C is not a K_C-subformula of A or B, we must have the following:

$$\mathsf{mc}{\downarrow}\frac{K_A\{\forall x P_A\} \vee K_B\{\forall x P_B\}}{K_C\left\{\mathsf{u}{\downarrow}\,\dfrac{\forall x\left(P_{C_1} \vee P_{C_2}\right)}{\forall x P_{C_1} \vee \exists x P_{C_2}}\right\}}$$

  we now have a number of cases to consider:

– If we have $P_A \equiv P_{C_1} \vee P_{C_2}$ and $P_B \equiv \mathsf{f}$ (WLOG), then we can construct:

$$\mathsf{mc}\!\downarrow \frac{K_A\left\{\mathsf{u}\!\downarrow \dfrac{\forall x\left(P_{C_1} \vee P_{C_2}\right)}{\forall x P_{C_1} \vee \exists x P_{C_2}}\right\} \vee K_B\left\{= \dfrac{\forall x \mathsf{f}}{\mathsf{f}}\right\}}{K_C\left\{\forall x P_{C_1} \vee \exists x P_{C_2}\right\}}$$

– If we have $P_A \equiv P_{C_1}$ and $P_B \equiv P_{C_2}$, then we can construct:

$$\mathsf{mc}\!\downarrow \frac{K_A\left\{\forall x P_{C_1}\right\} \vee K_B\left\{\begin{array}{c} \mathsf{n}\!\uparrow \dfrac{\forall x P_{C_2}}{[c \Rightarrow x]P_{C_2}} \\ \mathsf{n}\!\downarrow \dfrac{}{\exists x P_{C_2}} \end{array}\right\}}{K_C\left\{\forall x P_{C_1} \vee \exists x P_{C_2}\right\}}$$

– If we have $P_A \equiv P_{A_1} \vee P_{A_2}$ and $P_B \equiv P_{B_1} \vee P_{B_2}$, then we can construct:

$$\mathsf{mc}\!\downarrow \frac{K_A\left\{\mathsf{u}\!\downarrow \dfrac{\forall x\left(P_{A_1} \vee P_{A_2}\right)}{\forall x P_{A_1} \vee \exists x P_{A_2}}\right\} \vee K_B\left\{\mathsf{u}\!\downarrow \dfrac{\forall x\left(P_{B_1} \vee P_{B_2}\right)}{\forall x P_{B_1} \vee \exists x P_{B_2}}\right\}}{K_C\left\{\forall x P_{C_1} \vee \exists x P_{C_2}\right\}}$$

• If $\rho = \mathsf{r1}\!\downarrow$ then, since P_C is not a K_C-subformula of A or B, we must have the following:

$$\mathsf{mc}\!\downarrow \frac{K_A\{\forall x P_A\} \vee K_B\{\forall x P_B\}}{K_C\left\{\mathsf{u}\!\downarrow \dfrac{\forall x\left(P_{C_1} \vee P_{C_2}\right)}{\forall x P_{C_1} \vee P_{C_2}}\right\}}$$

we now have a number of cases to consider:

– If we have $P_A \equiv P_{C_1} \vee P_{C_2}$ and $P_B \equiv \mathsf{f}$ (WLOG), then we can construct:

$$\mathsf{mc}\!\downarrow \frac{K_A\left\{\mathsf{u}\!\downarrow \dfrac{\forall x\left(P_{C_1} \vee P_{C_2}\right)}{\forall x P_{C_1} \vee P_{C_2}}\right\} \vee K_B\left\{= \dfrac{\forall x \mathsf{f}}{\forall x \mathsf{f} \vee \mathsf{f}}\right\}}{K_C\left\{\forall x P_{C_1} \vee P_{C_2}\right\}}$$

– If we have $P_A \equiv P_{C_1}$ and $P_B \equiv P_{C_2}$, then P_{C_2} must be free for x, so we can construct:

$$\mathsf{mc}\!\downarrow \frac{K_A\left\{\forall x P_{C_1}\right\} \vee K_B\left\{\begin{array}{c} \mathsf{n}\!\uparrow \dfrac{\forall x P_{C_2}}{[c \Rightarrow x]P_{C_2}} \\ \mathsf{n}\!\downarrow \dfrac{}{\exists x P_{C_2}} \end{array}\right\}}{K_C\left\{\forall x P_{C_1} \vee \exists x P_{C_2}\right\}}$$

– If we have $P_A \equiv P_{A_1} \vee P_{A_2}$ and $P_B \equiv P_{B_1} \vee P_{B_2}$, then we can construct:

$$\mathsf{mc}\!\downarrow \frac{K_A\left\{\mathsf{u}\!\downarrow \dfrac{\forall x\left(P_{A_1} \vee P_{A_2}\right)}{\forall x P_{A_1} \vee \exists x P_{A_2}}\right\} \vee K_B\left\{\mathsf{u}\!\downarrow \dfrac{\forall x\left(P_{B_1} \vee P_{B_2}\right)}{\forall x P_{B_1} \vee \exists x P_{B_2}}\right\}}{K_C\left\{\forall x P_{C_1} \vee \exists x P_{C_2}\right\}}$$

- If $\rho = \mathsf{u}\!\uparrow$, we must have that the following:

$$\mathsf{mc}\!\downarrow \dfrac{K_A\left\{\left(P_{A_1} \wedge P_{A_2}\right)\right\} \vee K_B\left\{\left(P_{B_1} \wedge P_{B_2}\right)\right\}}{K_C\left\{\mathsf{u}\!\uparrow \dfrac{\forall x P_{C_1} \wedge \exists x P_{C_2}}{\exists x\left(P_{C_1} \wedge P_{C_2}\right)}\right\}}$$

  It must be the case that $P_{A_1} \equiv \mathsf{f}, P_{B_1} \equiv \mathsf{f}$ or both $P_{A_1} \equiv \forall x P'_{A_1}$ and $P_{B_1} \equiv \forall x P'_{B_1}$, with $P_{C_1} \in M_\vee(P'_{A_1}, P'_{B_1})$. Similarly, either $P_{A_2} \equiv \mathsf{f}, P_{B_2} \equiv \mathsf{f}$ or both $P_{A_2} \equiv \forall x P'_{A_2}$ and $P_{B_2} \equiv \forall x P'_{B_2}$, with $P_{C_2} \in M_\vee(P'_{A_2}, P'_{B_2})$. We will show two cases, the rest are similar:

  - If $P_{A_1} \equiv \mathsf{f}$, $P_{B_1} \equiv \forall x P_{C_1}$, $P_{A_2} \equiv \exists x P_{C_2}$ and $P_{B_2} \equiv \mathsf{f}$ then we can construct:

$$\mathsf{mc}\!\downarrow \dfrac{K_A\left\{\mathsf{u}\!\uparrow \dfrac{= \dfrac{\mathsf{f}}{\forall x \mathsf{f}} \wedge \exists x P_{C_2}}{\exists x\left(\mathsf{f} \wedge P_{C_2}\right)}\right\} \vee K_B\left\{\mathsf{u}\!\uparrow \dfrac{\forall x P_{C_1} \wedge = \dfrac{\mathsf{f}}{\exists x \mathsf{f}}}{\exists x\left(P_{C_1} \wedge \mathsf{f}\right)}\right\}}{K_C\left\{\exists x\left(P_{C_1} \wedge P_{C_2}\right)\right\}}$$

  - If $P_{A_1} \equiv \forall x P'_{A_1}$, $P_{B_1} \equiv \forall x P'_{B_1}$, $P_{A_2} \equiv \forall x P'_{A_2}$ and $P_{B_2} \equiv \forall x P'_{B_2}$ then we can construct:

$$\mathsf{mc}\!\downarrow \dfrac{K_A\left\{\mathsf{u}\!\uparrow \dfrac{\forall x P'_{A_1} \wedge \exists x P'_{A_2}}{\exists x\left(P'_{A_1} \wedge P'_{A_2}\right)}\right\} \vee K_B\left\{\mathsf{u}\!\uparrow \dfrac{\forall x P'_{B_1} \wedge \exists x P'_{B_2}}{\exists x\left(P'_{B_1} \wedge P'_{B_2}\right)}\right\}}{K_C\left\{\exists x\left(P_{C_1} \wedge P_{C_2}\right)\right\}}$$

Finally, we define two more reductions needed

$$r_3 : \quad \mathsf{u}\!\uparrow \dfrac{\mathsf{mc}\!\downarrow \dfrac{\exists x A \vee \exists x B}{\exists x C} \wedge \forall x D}{\exists x(C \wedge D)} \quad \longrightarrow \quad \mathsf{mc}\!\downarrow \dfrac{\mathsf{s}\!\uparrow \dfrac{(\exists x A \vee \exists x B) \wedge \mathsf{mc}\!\uparrow \dfrac{\forall x D}{\forall x D \wedge \forall x D}}{\mathsf{u}\!\uparrow \dfrac{\exists x A \wedge \forall x D}{\exists x(A \wedge D)} \vee \mathsf{u}\!\uparrow \dfrac{\exists x B \wedge \forall x D}{\exists x(B \wedge D)}}}{\exists x(C \vee D)}$$

$$r_4 : \quad \mathsf{u}\!\uparrow \dfrac{\mathsf{mc}\!\downarrow \dfrac{\forall x A \vee \forall x B}{\forall x C} \wedge \exists x D}{\exists x(C \wedge D)} \quad \longrightarrow \quad \mathsf{mc}\!\downarrow \dfrac{\mathsf{s}\!\uparrow \dfrac{(\forall x A \vee \forall x B) \wedge \mathsf{mc}\!\uparrow \dfrac{\exists x D}{\exists x D \wedge \exists x D}}{\mathsf{u}\!\uparrow \dfrac{\forall x A \wedge \exists x D}{\exists x(A \wedge D)} \vee \mathsf{u}\!\uparrow \dfrac{\forall x B \wedge \exists x D}{\exists x(B \wedge D)}}}{\exists x(C \vee D)}$$

Definition 3.42. We define the rewriting system for SKSmq,

$$\mathrm{M}_1 = \{r_1, r_2, r_3, r_4, s, t_1, t_2, t_3, m, \mathsf{mc}\!\downarrow - \mathsf{ai}\!\uparrow, \mathsf{mc}\!\uparrow - \mathsf{ai}\!\downarrow\}$$

.

As with M^+, we expect that M_1 is weakly normalising.

Conjecture 3.43. M_1 *is weakly normalising for derivations in* SKSmq. *Therefore every derivation can be decomposed in the following way:*

$$\begin{array}{c} A \\ \Vert \{\mathsf{mc}\uparrow\} \\ A' \\ \Vert \mathsf{SKSmq} \setminus \{\mathsf{mc}\downarrow, \mathsf{mc}\uparrow\} \\ B' \\ \Vert \{\mathsf{mc}\downarrow\} \\ B \end{array}$$

Chapter 4

Herbrand Proofs and Expansion Proofs

4.1 Herbrand Proofs

As discussed in the introduction, we will present two different conceptions of representing the "Herbrand content" of a proof: Herbrand proofs and expansion proofs. For each we define both a deep inference proof system—$\mathsf{KSh1}$ and $\mathsf{KSh2}$—and a class of proofs in each system that corresponds to Herbrand or expansion proofs, respectively. First, we will present $\mathsf{KSh1}$ and Herbrand proofs.

4.1.1 $\mathsf{KSh1}$ and Herbrand Proofs

As discussed in the introduction, Herbrand proofs consist of the following four steps:

1. Expansion of existential subformulae.
2. Prenexification/elimination of universal quantifiers.
3. Term assignment.
4. Propositional tautology check.

In [Brü06a], it is shown that all four of these steps can be carried out by inference rules in a deep inference system. To do so, we need to define a contraction rule that only operates on existential formulae.

Definition 4.1. We define the rule $\mathsf{qc}\downarrow$ to restrict contraction just to existential formulae:

$$\mathsf{qc}\downarrow \frac{\exists x A \vee \exists x A}{\exists x A}$$

Proposition 4.2. *$\mathsf{c}\downarrow$ is derivable for $\{\mathsf{ac}\downarrow, \mathsf{m}, \mathsf{qc}\downarrow, \mathsf{m}_2\downarrow\}$ and $\mathsf{qc}\downarrow$ is derivable for $\{\mathsf{ac}\downarrow, \mathsf{m}, \mathsf{m}_1\downarrow, \mathsf{m}_2\downarrow\}$.*

$$
\mathsf{qc}{\downarrow}\dfrac{\mathsf{r1}{\downarrow}\dfrac{\mathsf{r1}{\downarrow}\dfrac{=\dfrac{\mathsf{t}}{\forall y_1 \forall y_2 \boxed{=\dfrac{\mathsf{ai}{\downarrow}\dfrac{\mathsf{t}}{Py_1\vee\bar{P}y_1}\vee\left(\mathsf{aw}{\downarrow}\dfrac{\mathsf{f}}{\bar{P}c}\vee\mathsf{aw}{\downarrow}\dfrac{\mathsf{f}}{Py_2}\right)}{(\bar{P}c\vee Py_1)\vee(\bar{P}y_1\vee Py_2)}}}}{\forall y_1\boxed{(\bar{P}c\vee Py_1)\vee\mathsf{n}{\downarrow}\dfrac{\forall y_2(\bar{P}y_1\vee Py_2)}{\exists x_2\forall y_2(\bar{P}x_2\vee Py_2)}}}}{\mathsf{n}{\downarrow}\dfrac{\forall y_1(\bar{P}c\vee Py_1)}{\exists x_1\forall y_1(\bar{P}x_1\vee Py_1)}\vee\exists x_2\forall y_2(\bar{P}x_2\vee Py_2)}}{\exists x\forall y(\bar{P}x\vee Py)}
$$

Figure 4.1: A KSh1 proof of the "drinker's formula"

Proof. Straightforward. □

Definition 4.3. We define an alternate cut-free proof system for FOL, KSh1:

$$
\mathsf{KSh1} = \mathsf{KS} + \boxed{\begin{array}{ccc}
\mathsf{r1}{\downarrow}\dfrac{\forall x[A\vee B]}{[\forall xA\vee B]} & \mathsf{r2}{\downarrow}\dfrac{\forall x(A\wedge B)}{(\forall xA\wedge B)} & \mathsf{n}{\downarrow}\dfrac{[t\Rightarrow x]A}{\exists xA} \\[2ex]
\mathsf{r3}{\downarrow}\dfrac{\exists x[A\vee B]}{[\exists xA\vee B]} & \mathsf{r4}{\downarrow}\dfrac{\exists x(A\wedge B)}{(\exists xA\wedge B)} & \mathsf{qc}{\downarrow}\dfrac{\exists xA\vee\exists xA}{\exists xA}
\end{array}}
$$

with the same equality relation as SKSq. Following [Brü06a], we define a Herbrand proof in the context of KSh1 in the following way.

Definition 4.4. A closed KSh1 proof is a *Herbrand proof* if it is in the following form:

$$
\begin{array}{c}
\Vert\mathsf{KS} \\
\forall\vec{x}[\vec{t}\Rightarrow\vec{y}]B \\
\Vert\{\mathsf{n}{\downarrow}\} \\
Q\{B\} \\
\Vert\mathsf{RP}_{\downarrow} \\
A' \\
\Vert\{\mathsf{qc}{\downarrow}\} \\
A
\end{array}
$$

where $Q\{\ \}$ is a context consisting only of quantifiers and B is quantifier-free, and $\mathsf{RP}_{\downarrow}$ refers to the four "down" rules of passage.

Theorem 4.5 (Herbrand's Theorem for cut-free proofs [Brü06a])**.** *Let* $\begin{array}{c}\phi\Vert\mathsf{KSq}\\ A\end{array}$. *Then we can construct a Herbrand Proof of* A.

Remark 4.6. This proposition and proof are essentially the same proof of Herbrand's Theorem from a cut-free system in [Brü06a, Theorem 4.2]. However,

$$
\mathsf{qc}{\downarrow}\dfrac{\mathsf{r3}{\downarrow}\dfrac{\mathsf{n}{\downarrow}\dfrac{=\dfrac{\mathsf{t}}{\forall y_1 \forall y_2 \boxed{=\dfrac{\mathsf{ai}{\downarrow}\dfrac{\mathsf{t}}{Py_1 \vee \bar{P}y_1} \vee \left(\mathsf{aw}{\downarrow}\dfrac{\mathsf{f}}{\bar{P}c} \vee \mathsf{aw}{\downarrow}\dfrac{\mathsf{f}}{Py_2}\right)}{\left(\bar{P}c \vee Py_1\right) \vee \left(\bar{P}y_1 \vee Py_2\right)}}}}{\exists x_1 \boxed{\mathsf{r1}{\downarrow}\dfrac{\forall y_1 \boxed{\mathsf{r3}{\downarrow}\dfrac{\mathsf{n}{\downarrow}\dfrac{\left(\bar{P}x_1 \vee Py_1\right) \vee \left(\bar{P}y_1 \vee Py_2\right)}{\exists x_2 \boxed{\mathsf{r1}{\downarrow}\dfrac{\forall y_2\left(\left(\bar{P}x_1 \vee Py_1\right) \vee \left(\bar{P}x_2 \vee Py_2\right)\right)}{\left(\bar{P}x_1 \vee Py_1\right) \vee \forall y_2\left(\bar{P}x_2 \vee Py_2\right)}}}}{\left(\bar{P}x_1 \vee Py_1\right) \vee \exists x_2 \forall y_2\left(\bar{P}x_2 \vee Py_2\right)}}}{\forall y_1\left(\bar{P}x_1 \vee Py_1\right) \vee \exists x_2 \forall y_2\left(\bar{P}x_2 \vee Py_2\right)}}}}{\exists x_1 \forall y_1\left(\bar{P}x_1 \vee Py_1\right) \vee \exists x_2 \forall y_2\left(\bar{P}x_2 \vee Py_2\right)}}{\exists x \forall y\left(\bar{P}x \vee Py\right)}
$$

Figure 4.2: A Herbrand proof of the drinker's formula

the proof is worth reworking in the open deduction formalism, and also since its relation to cut elimination is different: in Brünnler's paper it is a corollary to cut elimination, whereas here we are using it to organise the part of the proof that does not contain first-order cuts into a useful form.

Before proving the theorem, we state and prove a crucial lemma.

Lemma 4.7. *We can carry out the following proof transformation:*

$$
\begin{array}{c} \phi \| \mathsf{KSq} \setminus \{\mathsf{m}_1{\downarrow}\} \\ A \end{array} \quad \longrightarrow \quad \begin{array}{c} \| \mathsf{KS} \cup \{\mathsf{n}{\downarrow}\} \\ Q\{A_P\} \\ \| \mathsf{RP}_{\downarrow} \\ A \end{array}
$$

where $Q\{\ \}$ is a sequence of quantifiers and A_P is the formula obtained by removing all quantifiers from A.

Proof. We proceed by induction on the length of ϕ, with the base case being trivial. If the final rule in ϕ is in KS, the inductive step is also trivial [Brü06a]. By Proposition 3.13, we can replace $\mathsf{u}{\downarrow}$ with $\{\mathsf{r1}{\downarrow}, \mathsf{n}{\downarrow}\}$, with $\mathsf{r1}{\downarrow}$ simply getting absorbed into $\mathsf{RP}_{\downarrow}$. Thus we are left with just $\mathsf{m}_2{\downarrow}$ and $\mathsf{n}{\downarrow}$ to deal with.

We can eliminate instances of $\mathsf{m}_2{\downarrow}$ in the following way:

$$
K\left\{\mathsf{m}_2{\downarrow}\dfrac{\begin{array}{c} \phi \| \mathsf{KSq} \setminus \{\mathsf{m}_1{\downarrow}\} \\ \forall x A \vee \forall y[y \Rightarrow x]B \end{array}}{\forall x(A \vee B)}\right\} \quad \xrightarrow{IH} \quad \begin{array}{c} \phi' \| \mathsf{KS} \cup \{\mathsf{n}{\downarrow}\} \\ Q_1\{\forall x Q_2\{\forall y Q_3\{A_P \vee [y \Rightarrow x]B_P\}\}\} \\ \| \mathsf{RP}_{\downarrow} \\ K\left\{\mathsf{m}_2{\downarrow}\dfrac{\forall x A \vee \forall y[y \Rightarrow x]B}{\forall x(A \vee B)}\right\} \end{array} \quad \longrightarrow
$$

$$
\begin{array}{c}
\phi' \Big\| \mathsf{KS}\cup\{\mathsf{n}\downarrow\} \\
Q_1\left\{\forall x Q_2\left\{\mathsf{n}\uparrow\dfrac{\forall y Q_3\{A_P \vee [y \Rightarrow x]B_P\}}{Q_3\{A_P \vee B_P\}}\right\}\right\} \\
\| \mathsf{RP}_\downarrow \\
K\{\forall x(A \vee B)\}
\end{array}
\xrightarrow{\text{Prop. 3.22}}
\begin{array}{c}
\phi' \Big\| \mathsf{KS}\cup\{\mathsf{n}\downarrow\} \\
Q_1\{\forall x Q_2\{Q_3\{A_P \vee B_P\}\}\} \\
\| \mathsf{RP}_\downarrow \\
K\{\forall x(A \vee B)\}
\end{array}
$$

Instances of $\mathsf{n}\downarrow$ are permuted above the rules of passage:

$$
K\left\{\mathsf{n}\downarrow\dfrac{\begin{array}{c}\phi \| \mathsf{KSq}\setminus\{\mathsf{m}_1\downarrow\} \\ [t \Rightarrow x]A\end{array}}{\exists x A}\right\}
\xrightarrow{IH}
\begin{array}{c}
\phi \Big\| \mathsf{KS}\cup\{\mathsf{n}\downarrow\} \\
Q_1\{Q_2\{[t \Rightarrow x]A_P\}\} \\
\| \mathsf{RP}_\downarrow \\
K\left\{\mathsf{n}\downarrow\dfrac{[t \Rightarrow x]A}{\exists x A}\right\}
\end{array}
\longrightarrow
\begin{array}{c}
\phi \Big\| \mathsf{KS}\cup\{\mathsf{n}\downarrow\} \\
Q_1\left\{\mathsf{n}\downarrow\dfrac{Q_2\{[t \Rightarrow x]A_P\}}{\exists x Q_2\{[t \Rightarrow x]A_P\}}\right\} \\
\| \mathsf{RP}_\downarrow \\
K\{\exists x A\}
\end{array}
$$

□

Proof of Theorem 4.5. We work in stages, creating one section of the Herbrand Proof at a time.

1. First of all, we refactorise the contractive rules, $\{\mathsf{ac}\downarrow, \mathsf{m}, \mathsf{qc}\downarrow, \mathsf{m}_2\downarrow\}$, into $\{\mathsf{ac}\downarrow, \mathsf{m}, \mathsf{m}_1\downarrow, \mathsf{m}_2\downarrow\}$. Then, we use the following rewrites to permute $\mathsf{qc}\downarrow$ down the proof:

$$
\mathsf{qc}\downarrow - \rho_1 : \quad
\rho\,\dfrac{K\left\{\mathsf{qc}\downarrow\dfrac{\exists x A \vee \exists x A}{\exists x A}\right\}}{K'\{\exists x A\}}
\longrightarrow
\rho\,\dfrac{K\{\exists x A \vee \exists x A\}}{K'\left\{\mathsf{qc}\downarrow\dfrac{\exists x A \vee \exists x A}{\exists x A}\right\}}
$$

$$
\mathsf{qc}\downarrow - \rho_2 : \quad
\mathsf{qc}\downarrow\dfrac{\exists x A \vee \exists x A}{\exists x \boxed{\rho\,\dfrac{A}{B}}}
\longrightarrow
\mathsf{qc}\downarrow\dfrac{\exists x \boxed{\rho\,\dfrac{A}{B}} \vee \exists x \boxed{\rho\,\dfrac{A}{B}}}{\exists x B}
$$

   Each of these reductions, if applied to the bottommost instance of $\mathsf{qc}\downarrow$, reduces the number of rules below the bottommost instance of $\mathsf{qc}\downarrow$.

2. By Lemma 4.7 we can now separate the remaining proof into a top half with $\mathsf{KS} \cup \{\mathsf{n}\downarrow\}$, and a bottom half consisting of $\mathsf{RP}_\downarrow$.

3. Since every other first-order rule is now eliminated from the proof, it is straightforward to permute $\mathsf{n}\downarrow$ rules down the proof.

Remark 4.8. The proof of Theorem 4.5 works exactly the same if KS is supplemented by an atomic cut rule, $\mathsf{ai}\uparrow$. Therefore, as Brünnler notes, only cuts with quantified eigenformulae need be eliminated to prove Herbrand's Theorem.

□

Proposition 4.9. *Given a Herbrand proof ϕ of A, we can construct a proof of A in the cut-free FOL system* KSq.

Proof. Immediate from Propositions 3.13, 3.15, 3.16, and 4.2. □

4.2 Expansion Proofs

4.2.1 Introduction

In [Mil87], Miller generalises the concept of the Herbrand expansion to higher order logic, representing the witness information in a tree structure, and explicit transformations between these ‘expansion proofs’ and cut-free sequent proofs are provided. Miller’s presentation of expansion proofs lacked some of the usual features of a formal proof system, crucially composition by an eliminable cut, but extensions in this direction have been carried out by multiple authors. In [Hei10], Heijltjes presents a system of ‘proof forests’, a graphical formalism of expansion proofs with cut and a local rewrite relation that performs cut elimination. Similar work has been carried out by McKinley [McK13] and more recently by Hetzl and Weller [HW13] and Alcolei et al. [Alc+17].

4.2.2 Expansion Trees

Remark 4.10. In this section, we will frequently use $\star$ in place of $\wedge$ and $\vee$, and Q in place of $\forall$ and $\exists$ if both cases can be combined into one. For clarity, we will sometimes distinguish between connectives in expansion trees, $\star_E$, and in formulae/derivations, $\star_F$.

Definition 4.11. We define *expansion trees*, the two functions *Sh* (shallow) and *Dp* (deep) from expansion trees to formulae, a set of *eigenvariables* $EV(E)$ for each expansion tree, and a partial function *Lab* from edges to terms, following [Mil87], [Hei10] and [CHM16]:

- Every literal A (including the units t and f) is an expansion tree. $Sh(A) :\equiv A$, $Dp(A) :\equiv A$, and $EV(A) = \emptyset$.

- If E_1 and E_2 are expansion trees with $EV(E_1) \cap EV(E_2) = \emptyset$, then $E_1 \star E_2$ is an expansion tree, with $Sh(E_1 \star_E E_2) :\equiv Sh(E_1) \star_F Sh(E_2)$, $Dp(E_1 \star_E E_2) :\equiv Dp(E_1) \star_F Dp(E_2)$, and $EV(E_1 \star E_2) = EV(E_1) \cup EV(E_2)$. We call $\star$ a *$\star$-node* and each unlabelled edge e_i connecting the $\star$-node to E_i a *$\star$-edge*. We represent $E_1 \star E_2$ as:

$$\begin{array}{ccc} E_1 & & E_2 \\ e_1\diagdown & & \diagup e_2 \\ & \star & \end{array}$$

- If E' is an expansion tree s.t. $Sh(E') \equiv A$ and $x \notin EV(E')$, then $E = \forall xA +^x E'$ is an expansion tree with $Sh(E) :\equiv \forall xA$, $Dp(E) :\equiv Dp(E')$, and $EV(E) := EV(E') \cup \{x\}$. We call $\forall xA$ a *$\forall$-node* and the edge e connecting the $\forall$-node and E' a *$\forall$-edge*, with $Lab(e) = x$. We represent E as:

$$\begin{array}{c} E' \\ e \,\big|\, x \\ \forall xA \end{array}$$

- If $t_1, \ldots, t_n$ are terms ($n \geq 0$), and $E_1, \ldots, E_n$ are expansion trees s.t. $x \notin EV(E_i)$ and $EV(E_i) \cap EV(E_j) = \emptyset$ for all $1 \leq i < j \leq n$, and $Sh(E_i) \equiv [t_i \Rightarrow x]A$, then $E = \exists xA +^{t_1} E_1 +^{t_2} \cdots +^{t_n} E_n$ is an expansion tree, where $Sh(E) :\equiv \exists xA$, $Dp(E) :\equiv Dp(E_1) \vee \cdots \vee Dp(E_n)$ (with $Dp(E) \equiv \mathsf{f}$ if $n = 0$), and $EV(E) = \bigcup_1^n EV(E_i)$. We call $\exists xA$ an $\exists$-*node* and each edge e_i connecting the $\exists$-node with E_i an $\exists$-*edge*, with $Lab(e_i) = t_i$. We represent E as:

$$\begin{tikzcd}[column sep=small]
E_1 \arrow[dr, dash, "e_1"', "t_1"] & \cdots & E_n \arrow[dl, dash, "e_n", "t_n"'] \\
& \exists xA &
\end{tikzcd}$$

Remark 4.12. Let ρ be a permutation of $[1 \ldots n]$. We consider the expansion trees $\exists xA +^{t_1} E_1 +^{t_2} \cdots +^{t_n} E_n$ and $\exists xA +^{t_{\rho(1)}} E_{\rho(1)} \cdots +^{t_{\rho(n)}} E_{\rho(n)}$ equal. Our trees are also presented the other way up to usual, e.g. [Hei10]. This is so that they are the same way up as the deep inference proofs we will translate them to below.

Definition 4.13. Let E be an expansion tree and let $<_E^-$ be the relation on the edges in E defined by:

- $e <_E^- e'$ if the node directly below e is the node directly above e'.

- $e <_E^- e'$ if e is an $\exists$-edge with $Lab(e) = t$, $x \in FV(t)$, e' is a $\forall$-edge and $Lab(e') = x$. In this case, we say e' *points to* e.

The *dependency relation* of E, $<_E$, is the transitive closure of $<_E^-$.

Definition 4.14. An expansion tree E is *correct* if $<_E$ is acyclic and $Dp(E)$ is a tautology. We can then call E an *expansion proof* of $Sh(E)$.

Example 4.15. Below is an expansion tree E, with $Sh(E) \equiv \exists x \forall y \big(\bar{P}x \vee Py\big)$ and $Dp(E) \equiv \big(\bar{P}c \vee Py_1\big) \vee \big(\bar{P}y_1 \vee Py_2\big)$. The tree is presented with all edges explicitly named, to define the dependency relation below, as well as the labels for the $\exists$-edges and $\forall$-edges.

$$\begin{tikzcd}[column sep=tiny]
\bar{P}c \arrow[dr, dash, "e_1"'] & & Py_1 \arrow[dl, dash, "e_2"] & & \bar{P}y_1 \arrow[dr, dash, "e_3"'] & & Py_2 \arrow[dl, dash, "e_4"] \\
& \vee \arrow[d, dash, "e_5"', "y_1"] & & & & \vee \arrow[d, dash, "y_2"', "e_6"] & \\
& \forall y_1[\bar{P}c \vee Py_1] \arrow[drr, dash, "e_7"', "c"] & & & & \forall y_2[\bar{P}y_1 \vee Py_2] \arrow[dll, dash, "y_1", "e_8"'] & \\
& & & \exists x \forall y[\bar{P}x \vee Py] & & &
\end{tikzcd}$$

The dependency relation is generated by the following inequalities: $e_3, e_4 < e_6 < e_8$ and $e_1, e_2 < e_5 < e_7$ and $e_8 < e_5$. e_5 points to e_8. As this dependency relation is acyclic and $\big(\bar{P}c \vee Py_1\big) \vee \big(\bar{P}y_1 \vee Py_2\big)$ is a tautology, E is correct, and thus an expansion proof.

4.2.3 Expansion Proofs with Cut

Definition 4.16. If $E, E_1, \ldots, E_n$ are expansion trees, with $Sh(E_i) \equiv \exists x A_i \wedge \forall x \bar{A}_i$ for each E_i, then $EC = E + \perp(E_1, \ldots, E_n)$ is an *expansion tree with cut*. We call $\perp$ the *cut node*, and each edge e_i connecting $\perp$ and E_i a *cut edge*. If $FV(\exists x A_i) = \emptyset$, then we say e_i is a *closed* cut edge. We represent EC as:

$$\mathrm{E} + \begin{tikzcd} E_1 \arrow[dr, dash, "e_1"'] & & E_n \arrow[dl, dash, "e_n"] \\ & \perp & \end{tikzcd}$$

We extend the deep and shallow functions to expansion trees with cuts: $Sh(E + \perp(E_1, \ldots, E_n)) \equiv A$, $Dp(E + \perp(E_1, \ldots, E_n)) \equiv Dp(E) \vee Dp(E_1) \vee \ldots \vee Dp(E_n)$.

Extending the notion of correctness to expansion trees with cut is straightforward.

Definition 4.17. The dependency relation for expansion trees with cut is the same as that for expansion trees, with the following addition to the definition of $<_E^-$:

- $e <_E^- e'$ if e is a cut edge connecting $\perp$ and E_i, $Sh(E_i) \equiv \forall y A_i \wedge \exists y \bar{A}_i$, $x \in FV(A_i)$, e' is a $\forall$-edge and $Lab(e') = x$. We still say that *e' points to e*.

The correctness criteria for expansion trees with cuts are the same for expansion trees, giving us *expansion proofs with cut*.

Remark 4.18. If every cut edge is closed for an expansion tree with cut, then the correctness criteria is exactly the same as for the expansion tree obtained by replacing the cut node with a series of $\wedge$ nodes.

At this point, we are close to being able to borrow the cut elimination method from Heijltjes' *Proof Forests* formalism [Hei10]. However, proof forests are only a subclass of what we define as expansion trees here. Therefore it will be useful to properly define this subclass, as well as the class of expansion proof with closed cuts, which will be another useful subclass later on. However, there is one further difference between our expansion proofs and the proof forests of Heijltjes. While the graphical structure of his (prenex) expansion trees is the same as ours, the correctness condition is different, defined in terms of the propositional validity of *switchings* of expansion trees. Since, however, this correctness criterion is equivalent to ours on prenex expansion trees, we need not define this different criterion in detail.

Definition 4.19. A *prenex* expansion tree is an expansion tree where no Q-node is above a $\star$-node.

If $E = E_1 \vee \ldots \vee E_n$ with E_i prenex expansion trees, then E is a *forest-style expansion tree*. If E is correct it is a *forest-style expansion proof.*

If $E = E' + \perp(E_1 \wedge F_1, \ldots, E_n \wedge F_n)$ with E' a forest-style expansion tree, and $E_1, F_1, \ldots, E_n, F_n$ are prenex expansion trees, then E is a *forest-style expansion tree with cut*. If F_c is correct, then it is a *forest-style expansion proof with cut*.

If E is an expansion tree with cut with every cut edge closed, we say that E is a *expansion tree with closed cut*. If F_c is correct, then it is a *expansion proof with closed cut*.

Convention 4.20. We consider every expansion tree/proof to also be an expansion tree/proof with cut.

Theorem 4.21. *Let $E = E' + \bot(E_1 \wedge F_1, \ldots, E_n \wedge F_n)$ be an expansion tree with cut, with $Sh(F_c) \equiv A$ and where E', E_i and F_i are all correct. Then we can construct from it a cut-free expansion proof $E_F = E_{F_1} \vee \ldots \vee E_{F_n}$ where E_{F_i} are prenex expansion trees and $Sh(E_F) \equiv A$.*

Proof. [Hei10, Proposition 16 and Theorem 21] □

4.2.4 KSh2 and Herbrand Normal Form

To aid the translation between open deduction proofs and expansion proofs, we introduce a slightly different cut-free proof system to KSh1. It involves two new rules.

Definition 4.22. We define the rule h↓, which we call a *Herbrand expander* and the rule ∃w↓, which we call *existential weakening*:

$$\mathsf{h}\!\downarrow \frac{\exists x A \vee [t \Rightarrow x]A}{\exists x A} \qquad \exists\mathsf{w}\!\downarrow \frac{\mathsf{f}}{\exists x A}$$

For technical reasons again, we insist that $[t \Rightarrow x]A$ is in fact $[t \Rightarrow x]A'$, where A' is an α-equivalent formula to A with fresh variables for all quantifiers, but for simplicity we will usually denote it A.

Remark 4.23. Unlike the n↓ rule, the h↓ Herbrand expander rule is invertible. Similar rules have been used in first-order sequent calculus systems for automated reasoning, such as Kanger's LC [DV01; Kan57] and also in sequent systems for translation to expansion proofs [AHW18].

Definition 4.24. We define the alternate cut-free system for FOL KSh2:

$$\mathsf{KSh2} = \mathsf{KS} + \boxed{\begin{array}{cc} \mathsf{r1}\!\downarrow \dfrac{\forall x[A \vee B]}{[\forall x A \vee B]} & \mathsf{h}\!\downarrow \dfrac{\exists x A \vee [t \Rightarrow x]A}{\exists x A} \\ \mathsf{r2}\!\downarrow \dfrac{\forall x(A \wedge B)}{(\forall x A \wedge B)} & \exists\mathsf{w}\!\downarrow \dfrac{\mathsf{f}}{\exists x A} \end{array}} + \boxed{\begin{array}{rclrcl} \forall x A &=& \forall z\{z \Rightarrow x\}A & \exists z A &=& \exists z[z \Rightarrow x]A \\ \forall x \forall y A &=& \forall y \forall x A & \exists x \exists y A &=& \exists y \exists x A \\ \forall x \mathsf{t} &=& \mathsf{t} = \exists x \mathsf{t} & \forall x \mathsf{f} &=& \mathsf{f} = \exists x \mathsf{f} \end{array}}$$

Where z does not occur in A for the top two equalities.

Remark 4.25. The ∃w↓ rule is derivable for KSh2\{∃w↓}, but we explicitly include it so that we can restrict weakening instances in certain parts of proofs.

$$
\mathsf{r1}\!\downarrow \frac{
=\frac{
=\frac{\mathsf{t}}{\forall y_1 \forall y_2 \boxed{\mathsf{ai}\!\downarrow \frac{\mathsf{t}}{Py_1 \vee \bar{P}y_1} \vee \left(\mathsf{aw}\!\downarrow \frac{\mathsf{f}}{\bar{P}c} \vee \mathsf{aw}\!\downarrow \frac{\mathsf{f}}{Py_2}\right)}}
}{
\forall y_1 \boxed{\mathsf{r1}\!\downarrow \frac{\forall y_2 \boxed{\left(\mathsf{w}\!\downarrow \frac{\mathsf{f}}{\exists x \forall y (\bar{P}x \vee Py)} \vee (\bar{P}y_1 \vee Py_2)\right) \vee (\bar{P}c \vee Py_1)}}{\mathsf{r1}\!\downarrow \frac{\forall y_2 \big(\exists x \forall y (\bar{P}x \vee Py) \vee (\bar{P}y_1 \vee Py_2)\big)}{\mathsf{h}\!\downarrow \frac{\exists x \forall y (\bar{P}x \vee Py) \vee \forall y_2 (\bar{P}y_1 \vee Py_2)}{\exists x \forall y (\bar{P}x \vee Py)}} \vee (\bar{P}c \vee Py_1)}}
}
}{
\mathsf{h}\!\downarrow \frac{\exists x \forall y (\bar{P}x \vee Py) \vee \forall y_1 (\bar{P}c \vee Py_1)}{\exists x \forall y (\bar{P}x \vee Py)}
}
$$

Figure 4.3: A proof of the drinker's formula in HNF

Definition 4.26. If ϕ is a closed $\mathsf{KSh2}$ proof in the following form, where $\forall\vec{x}$ is a list of universal quantifiers with distinct variables, and $Lo(\phi)$ is regular and in sequential form, we say ϕ is in *Herbrand Normal Form* (HNF):

$$
\begin{array}{c}
Up(\phi) \,\| \,\mathsf{KS} \\
\forall \vec{x} H_\phi(A) \\
\| \, \{\exists \mathsf{w}\!\downarrow\} \\
\forall \vec{x} H^+_\phi(A) \\
Lo(\phi) \,\| \, \{\mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow, \mathsf{h}\!\downarrow\} \\
A
\end{array}
$$

$H_\phi(A)$, the *Herbrand disjunction of A according to ϕ*, or just the *Herbrand disjunction of A*, contains no quantifiers, whereas $H^+_\phi(A)$, the *expansive Herbrand disjunction of A according to ϕ*, may contain quantifiers. $Up(\phi)$ is called the *upper part* of ϕ, and $Lo(\phi)$ the *lower part* of ϕ.

We also define proofs in HNF with cut. Notice that the cuts do not necessarily need to be at the bottom of the proof, and need not be closed.

Definition 4.27. We define $\mathsf{KSh2c} = \mathsf{KSh2} \cup \{\mathsf{qi}\!\uparrow\}$, an alternate first-order system with cut.

Definition 4.28. If ϕ is a closed $\mathsf{KSh2c}$ proof in the following form, then we say ϕ is in *Herbrand Normal Form with Cut* (HNFC)

$$
\begin{array}{c}
Up(\phi) \,\| \,\mathsf{KS} \\
\forall \vec{x} H_\phi(A) \\
\| \, \{\exists \mathsf{w}\!\downarrow\} \\
\forall \vec{x} H^+_\phi(A) \\
Lo(\phi) \,\| \, \{\mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow, \mathsf{h}\!\downarrow, \mathsf{qi}\!\uparrow\} \\
A
\end{array}
$$

In this thesis, we will only ever translate between proofs in HNF and Herbrand proofs when they are cut-free or with closed cuts. Therefore a cut free translation between the two forms suffices.

Proposition 4.29. *A formula A has a proof in HNF (*$\mathsf{KSh2}$*) if and only if it has a Herbrand proof (*$\mathsf{KSh1}$*).*

Proof. Let ϕ be a proof of A in HNF. As $H_\phi(A)$ is the Herbrand expansion of A, it is straightforward to construct a Herbrand proof for A: one can infer the necessary $\mathsf{n}\!\downarrow$ and $\mathsf{qc}\!\downarrow$ rules by comparing $H_\phi(A)$ and A. Now let ϕ be a Herbrand Proof. The order of the quantifiers in $Q\{\ \}$ (as in Definition 4.4) is used to build the HNF proof. Thus, we proceed by induction on the number of quantifiers in $Q\{\ \}$. If there are none, it is obviously trivial. We split the inductive step into two cases.

First, consider ϕ_1 of the form shown, where P is a quantifier-free context and $Q\{\ \} = \forall z Q'\{\ \}$. Clearly ϕ_2 is also a Herbrand proof, so by the IH the proof ϕ_3 in HNF is constructible, from which we can construct ϕ_4.

$$
\begin{array}{cccc}
\begin{array}{c}
\Vert\,\mathsf{KS} \\
\forall z \forall \vec{x} B\{\vec{y} \Leftarrow \vec{t}\} \\
\Vert\,\{\mathsf{n}\!\downarrow\} \\
\forall z Q'\{B\} \\
\Vert\,\mathsf{RP}_\downarrow \\
P\{\forall z C'\} \\
\Vert\,\{\mathsf{qc}\!\downarrow\} \\
P\{\forall z C\}
\end{array}
&
\begin{array}{c}
\Vert\,\mathsf{KS} \\
\forall \vec{x} B\{\vec{y} \Leftarrow \vec{t}\} \\
\Vert\,\{\mathsf{n}\!\downarrow\} \\
Q'\{B\} \\
\Vert\,\mathsf{RP}_\downarrow \\
P\{C'\} \\
\Vert\,\{\mathsf{qc}\!\downarrow\} \\
P\{C\}
\end{array}
&
\begin{array}{c}
Up\phi_3 \Vert\,\mathsf{KS} \\
\forall \vec{x} H_{\phi_3} P\{C\} \\
\Vert\,\{\exists\mathsf{w}\!\downarrow\} \\
\forall \vec{x} H^+_{\phi_3} P\{C\} \\
Lo(\phi_3) \Vert\,\{\mathsf{r1}\!\downarrow,\mathsf{r2}\!\downarrow,\mathsf{h}\!\downarrow\} \\
P\{C\}
\end{array}
&
\begin{array}{c}
\forall z Up\phi_3 \Vert\,\mathsf{KS} \\
\forall z \forall \vec{x} H_{\phi_3} P\{C\} \\
\Vert\,\{\exists\mathsf{w}\!\downarrow\} \\
\forall z \forall \vec{x} H^+_{\phi_3} P\{C\} \\
\forall z Lo\phi_3 \Vert\,\{\mathsf{r1}\!\downarrow,\mathsf{r2}\!\downarrow,\mathsf{h}\!\downarrow\} \\
\forall z P\{C\} \\
\Vert\,\{\mathsf{r1}\!\downarrow,\mathsf{r2}\!\downarrow\} \\
P\{\forall z C\}
\end{array}
\\
\phi_1 & \phi_2 & \phi_3 & \phi_4
\end{array}
$$

In the same way, we consider the case where $Q\{\ \} = \exists z Q'\{\ \}$. Below we only show the case where there is no contraction acting on $\exists z C$, but the case with such a contraction is similar.

$$
\begin{array}{cccc}
\begin{array}{c}
\Vert\,\mathsf{KS} \\
\forall \vec{x} B\{\vec{y} \Leftarrow \vec{t}\}\{z \Leftarrow t\} \\
\Vert\,\{\mathsf{n}\!\downarrow\} \\
\exists z Q'\{B\} \\
\Vert\,\mathsf{RP}_\downarrow \\
P\{\exists z C'\} \\
\Vert\,\{\mathsf{qc}\!\downarrow\} \\
P\{\exists z C\}
\end{array}
&
\begin{array}{c}
\Vert\,\mathsf{KS} \\
\forall \vec{x} B\{\vec{y} \Leftarrow \vec{t}\}\{z \Leftarrow t\} \\
\Vert\,\{\mathsf{n}\!\downarrow\} \\
Q'\{B\}\{z \Leftarrow t\} \\
\Vert\,\mathsf{RP}_\downarrow \\
P\{C'\{z \Leftarrow t\}\} \\
\Vert\,\{\mathsf{qc}\!\downarrow\} \\
P\{C\{z \Leftarrow t\}\}
\end{array}
&
\begin{array}{c}
Up(\phi_3) \Vert\,\mathsf{KS} \\
\forall \vec{x} P\{D\{z \Leftarrow t\}\} \\
\Vert\,\{\exists\mathsf{w}\!\downarrow\} \\
\forall \vec{x} P\{D^+\{z \Leftarrow t\}\} \\
Lo(\phi_3) \Vert\,\{\mathsf{r1}\!\downarrow,\mathsf{r2}\!\downarrow,\mathsf{h}\!\downarrow\} \\
P\{C\{z \Leftarrow t\}\}
\end{array}
&
\begin{array}{c}
Up(\phi_3) \Vert\,\mathsf{KS} \\
\forall \vec{x} P\{D\{z \Leftarrow t\}\} \\
\Vert\,\{\exists\mathsf{w}\!\downarrow\} \\
\forall \vec{x} P\{\exists z C \vee D^+\{z \Leftarrow t\}\} \\
Lo(\phi_3) \Vert\,\{\mathsf{r1}\!\downarrow,\mathsf{r2}\!\downarrow,\mathsf{h}\!\downarrow\} \\
P\left\{ \mathsf{h}\!\downarrow \dfrac{\exists z C \vee C\{z \Leftarrow t\}}{\exists z C} \right\}
\end{array}
\\
\phi_1 & \phi_2 & \phi_3 & \phi_4
\end{array}
$$

where $P\{D\{z \Leftarrow t\}\} \equiv H_{\phi_3}(P\{C\{z \Leftarrow t\}\})$ and $P\{D^+\{z \Leftarrow t\}\} \equiv H^+_{\phi_3}(P\{C\{z \Leftarrow t\}\})$. $\square$

4.3 Translations Between Proofs in HNF and Expansion Proofs

Above, we gave translations between Herbrand proofs in KSh1 and KSh2 proofs in HNF. We will now give a translations between KSh2c proofs in HNFC and expansion proofs with cut, thus giving us a link between deep inference Herbrand proofs and expansion proofs. In the paper [Ral18], we showed translations between Herbrand proofs and cut-free expansion proofs. Here, we will show translations between Herbrand proofs with cut and expansion proofs with cut, but these will be conservative extensions of the cut-free translations, so we use the same terminology.

4.3.1 HNFC to Expansion Proofs

Before stating and proving the main theorem, we will define the map π_1 from KSh2c proofs to expansion proofs with cut (from now on in this section we will often omit "with cut" if unambiguous), and then prove some lemmas to help prove that the dependency relation in all expansion proofs in the range of π_1 is acyclic.

Remark 4.30. We extend the notion and syntax of contexts from derivations to expansion trees with cut. For this to make sense, a context can only take expansion trees with the same shallow formula.

Definition 4.31. We define a map π_1' from the lower part of KSh2c proofs in HNFC to expansion trees in the following way, working from the bottom.

On the conclusion of ϕ, we define π_1' as follows:

- $\pi_1'(B \star C) = \pi_1'(B) \star \pi_1'(C)$
- $\pi_1'(\forall x B) = \forall x B +^x \pi_1'(B)$
- $\pi_1'(\exists x B) = \exists x B$

The r1$\downarrow$ and r2$\downarrow$ rules are ignored by expansion trees, each h$\downarrow$ rule adds a branch to a $\exists$-node, and each qi$\downarrow$ rule adds another cut edge:

- If $\phi \equiv \begin{array}{c} K\left\{ \mathsf{r1}\!\downarrow \dfrac{\forall x(B \vee C)}{\forall x B \vee C} \right\} \\ \phi' \| \\ A \end{array}$ then $\pi_1'(\phi) = \pi_1' \left(\begin{array}{c} K\{\forall x B \vee C\} \\ \phi' \| \\ A \end{array} \right)$.
- If $\phi \equiv \begin{array}{c} K\left\{ \mathsf{r2}\!\downarrow \dfrac{\forall x(B \wedge C)}{\forall x B \wedge C} \right\} \\ \phi' \| \\ A \end{array}$ then $\pi_1'(\phi) = \pi_1' \left(\begin{array}{c} K\{\forall x B \wedge C\} \\ \phi' \| \\ A \end{array} \right)$.

- If $\pi_1'\left(\begin{array}{c} K\{\exists xB\} \\ \phi\| \\ A \end{array}\right) = K_{\pi_1}(\exists xB +^{\tau_1} E_1 + \cdots +^{\tau_n} E_n)$, then:

$$\pi_1'\left(\begin{array}{c} K\left\{ \mathsf{h}\!\downarrow \dfrac{\exists xB \vee [\tau_{n+1} \Rightarrow x]B}{\exists xB} \right\} \\ \phi\| \\ A \end{array}\right) = K_{\pi_1}(\exists xB +^{\tau_1} E_1 + \cdots +^{\tau_{n+1}} E_{n+1})$$

where $E_{n+1} = \pi_1'([\tau_{n+1} \Rightarrow x]B)$.

- If $\pi_1'\left(\begin{array}{c} K\{\mathsf{f}\} \\ \phi\| \\ A \end{array}\right) = E + \bot(E_1, \ldots, E_n)$, then:

$$\pi_1'\left(\begin{array}{c} K\left\{ \mathsf{qi}\!\uparrow \dfrac{\exists xB \wedge \forall xB}{\mathsf{f}} \right\} \\ \phi\| \\ A \end{array}\right) = E + \bot(E_1, \ldots, E_n, \pi_1'(\exists xB \wedge \forall xB))$$

We then define the map π_1 from KSh2 proofs in HNF to expansion trees as $\pi_1(\phi) = \pi_1'(Lo(\phi))$.

To show that $\pi_1(\phi)$ is an expansion proof, we need to prove that $\forall\vec{x}H_\phi(A)$ is a tautology and $<_E$ is acyclic. As $\forall\vec{x}H_\phi(A)$ has a proof in KS it is a tautology. Thus all that is needed is the acyclicity of $<_E$. To do so, we define the following partial order on variables in the lower part of KSh2c proofs in HNFC.

Definition 4.32. Let ϕ be a proof in HNFC. Define the partial order $<_\phi$ on the variables of occurring in $Lo(\phi)$ to be the minimal partial order such that $y <_\phi x$ if $K_1\{Q_1xK_2\{Q_2yB\}\}$ is a section of $Lo(\phi)$.

Proposition 4.33. *$<_\phi$ is well-defined for all* KSh2c *proofs in HNFC.*

Proof. Let ϕ be a proof of A in HNF, as in Definition 4.26. As $Lo(\phi)$ only contains $\mathsf{h}\!\downarrow, \mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow$ and $\mathsf{qi}\!\uparrow$ rules and no α-substitution, if a variable v occurs in $Lo(\phi)$ then v occurs in $\forall xH^+_\phi(A)$. Notice also that none of $\mathsf{h}\!\downarrow, \mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow$ or $\mathsf{qi}\!\uparrow$ can play the role of ρ in the following scheme:

$$\rho \frac{K\{Q_1v_1A_1\}\{Q_2v_2A_2\}}{K_1\{Q_1v_1\{K_2Q_2v_2B\}\}} \quad .$$

Therefore, if $K_1\{Q_1xK_2\{Q_2yB\}\}$ is a section of $Lo(\phi)$, then $\forall xH^+_\phi(A)$ is of the form $K_3\{Q_1xK_4\{Q_2yC\}\}$, i.e. no dependencies can be introduced below $\forall xH^+_\phi(A)$. Thus $x <_\phi y$ iff. $\forall xH^+_\phi(A)$ can be written $L_1\{Q_1xL_2\{Q_2yC\}\}$ for some $L_1\{\ \}, L_2\{\ \}$ and C and is therefore a well-defined partial order. $\square$

We now need to make sure the partial order $<_\phi$ works in the same way as the dependency relation on the expansion proof: that pointing and the descendant relation induces the correct ordering on variables.

Lemma 4.34. *Let ϕ be an* KSh2c *proof in HNFC and e' an $\forall$-edge in $\pi_1(\phi)$ that points to the $\exists$-edge e. If $Lab(e') = y$ and the $\exists$-node below e is $\exists xA$, then $x <_\phi y$.*

Proof. Since we have an $\exists$-node $\exists xA$ in $\pi_1(\phi)$ with an edge labelled t below it, there must be the following $\mathsf{h}\!\downarrow$ rule in ϕ:

$$K\left\{\mathsf{h}\!\downarrow\frac{\exists xA \vee [\tau \Rightarrow x]A}{\exists xA}\right\}$$

Since e points to e', y must occur freely in t. As ϕ is closed, y cannot be a free variable in $K\{\exists xA \vee [\tau \Rightarrow x]A\}$. Thus $K\{\ \}$ must be of the form $K_1\{\forall y K_2\{\ \}\}$. Therefore $x <_\phi y$. □

Lemma 4.35. *Let ϕ be an* KSh2c *proof in HNFC and e' an $\forall$-edge in $\pi_1(\phi)$ that points to the cut-edge e. If $Lab(e') = y$, E is the expansion tree below the cut edge e with $Sh(E) \equiv A \wedge \bar{A}$ and Qx is some quantifier appearing in A (with $\bar{Q}x$ appearing in $\bar{A}$), then $x <_\phi y$.*

Proof. The cut-edge e in $\pi_1(\phi)$ corresponds to some cut $K\left\{\mathsf{qi}\!\uparrow\frac{A \wedge \bar{A}}{\mathsf{f}}\right\}$ in ϕ. Since e' points to e, we know that $y \in FV(A)$. But we also know that ϕ is a closed proof. Therefore $K\{\ \} = K_1\{\forall y K_2\{\ \}\}$, and $x <_\phi y$. □

Lemma 4.36. *Let ϕ be an* KSh2c *proof in HNFC, e a $\forall$-edge of $\pi_1(\phi)$ labelled by x and e' an $\exists$-edge above an $\exists$-node $\exists yA$. If e is a descendant of e' then $x <_\phi y$.*

Proof. $Sh(\pi_1(\phi)) \equiv K_1\{\exists y K_2 \forall x\{B\}\}$ (for some $K_1\{\ \}, K_2\{\ \}$, and B) is the conclusion of ϕ, so $x <_\phi y$. □

Lemma 4.37. *Let ϕ be an* KSh2c *proof in HNFC, $E_\phi = \pi_1(\phi)$ and e and e' be (not necessarily distinct) $\forall$-edges in E_ϕ s.t. $e <_{E_\phi} e'$, $Lab(e) = x$ and $Lab(e') = x'$. Then $x <_\phi x'$.*

Proof. As $e <_{E_\phi} e'$, there must be a chain

$$e_{q_0} <^-_{E_\phi} \cdots <^-_{E_\phi} e_{p_1} <^-_{E_\phi} e_{q_1} <^-_{E_\phi} \cdots <^-_{E_\phi} e_{p_m} <^-_{E_\phi} e_{q_m} <^-_{E_\phi} \cdots <^-_{E_\phi} e_{p_n}$$

where $e_{q_0} = e$ and $e_{p_n} = e'$, e_{q_i} points to e_{p_i}, and e_{q_i} is a descendant of $e_{p_{i+1}}$ in the expansion tree. If e_{q_i} points to e_{p_i}, then either e_{p_i} is an $\exists$-node or a cut node.

If e_{p_i} is an $\exists$-node, then by Lemma 4.34, we know that if $\exists x_{p_i}$ is the node above p_i and $Lab(e_{q_i}) = x_{q_i}$, then $x_{p_i} <_\phi x_{q_i}$. By Lemma 4.36, since $e_{q_{i-1}}$ is a descendant of e_{p_i} in the expansion tree, $x_{q_{i-1}} <_\phi x_{p_i}$, so we have $x_{q_{i-1}} <_\phi x_{q_i}$.

If e_{p_i} is a cut node, then we know that, since $e_{q_{i-1}}$ is a descendent of e_{p_i}, by Lemma 4.35, $x_{q_{i-1}} <_\phi x_{q_i}$.

Therefore, we have that $e_{q_0} <_\phi e_{q_{n-1}}$. Since $e_{q_{n-1}}$ must be a descendent of e_{p_n}, we have that $x = e_{q_0} <_\phi e_{p_n} = x'$. □

We are now ready to prove the main theorem of this subsection: using our defined partial order to show that the dependency relation on the expansion tree is acyclic.

Theorem 4.38. *Let ϕ be a* KSh2c *proof of A in HNFC. Then we can construct an expansion proof with cut $E_\phi = \pi_1(\phi)$, with $Sh(E_\phi) \equiv A$, and $Dp(E_\phi) \equiv H_\phi(A)$.*

Proof. As described above, we only need to show that the dependency relation of E_ϕ is acyclic. Assume there were a cycle in $<_{E_\phi}$. Clearly, it could not be generated by just by travelling up the expansion tree. Thus, there is some $\forall$-edge e and an $\exists$-edge or cut edge e' such that e points to e' and $e <_{E_\phi} e' <_{E_\phi} e$. But then, if $Lab(e) = x$, by Lemma 4.37, $x <_\phi x$. But this contradicts Proposition 4.33. Therefore $<_{E_\phi}$ is acyclic. □

4.3.2 Expansion Proofs to HNFC

For the translation from expansion proofs to KSh2c proofs in HNFC, we show that we can progressively build up a KSh2c proof by working through the "minimal" nodes of an expansion proof. Unlike the previous translation, there is not necessarily a unique proof corresponding to each expansion proof, but a total order on universally quantified variables that respects $<_E$ is sufficient to give a unique proof up to equalities.

Convention 4.39. We will not tend to omit "with cut" in this section, as there are a few points where the distinction between cut-free and cut-full expansion proofs is important.

First, we need to relax the definition of expansion trees in a few ways, so we have objects that can correspond to incomplete HNF proofs.

Definition 4.40. A *weak* expansion tree is defined in the same way as in Definition 4.11 except that the first condition is weakened to allow any formula to be a leaf of the tree. A weak expansion tree with an acyclic dependency relation is correct regardless of whether its deep formula is a tautology.

A *weak* expansion tree with cut $E + \bot(E_1, \ldots, E_n)$ is just an expansion tree with cut where E and E_i are allowed to be weak expansion trees.

Definition 4.41. We define the *expansive deep formula $Dp^+(E)$* for (weak) expansion trees, which is defined in the same way as the usual deep formula except that:

$$Dp^+(\exists xA +^{t_1} E_1 +^{t_2} \cdots +^{t_n} E_n) :\equiv \exists xA \vee Dp^+(E_1) \vee \ldots \vee Dp^+(E_n)$$

Now, we define minimal edges and nodes of expansion trees, so that we have a strategy to translate the tree node by node.

Definition 4.42. A *minimal edge* of a (weak) expansion tree (with cut) E is an edge that is minimal w.r.t. to $<_E$.

If all the edges below a node are minimal, we say that the node is a *minimal node*.

Lemma 4.43. *If E is a weak expansion proof with cut with no minimal edges below existential nodes and no minimal universal and cut nodes, then it has a minimal $\star$ node.*

Proof. Assume E is a weak expansion proof with no minimal edges below existential nodes and no minimal universal or cut nodes. Clearly, there must be must be at least one minimal edge e_0, and by the assumption it must be below a node $\star_0$. Let e'_0 be the other edge below $\star_0$. If e'_0 is minimal, we are done. If not, pick some minimal edge $e_1 < e'_0$, which again, with $e'_1 < e'_0$, must be below some $\star_1$. For each e'_i that is not minimal, we can find $e'_{i+1} < e'_i$. As E is finite, this sequence cannot continue indefinitely, so eventually we will find two minimal edges e_n and e'_n below $\star_n$. □

Now, we show that we can delete minimal nodes of (weak) expansion trees, converting the deleted information to a part of an HNF proof.

Lemma 4.44. *Let $E = K_E\{\forall xA +^x A\}$, with $Dp^+(E) \equiv K\{A\}$, be a correct weak expansion tree with a minimal $\forall$-edge labelled by x (which we will call e). Then there is a derivation* $\begin{array}{c}\forall xK\{A\}\\ \Vert\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\}\\ K\{\forall xA\}\end{array}$.

Proof. We proceed by induction on the height of the node $\forall xA$ in E. If $\forall xA$ is the bottom node, then $K\{A\} \equiv A$ and we are done. Let E be an expansion tree where $\forall x$ is not the bottom node. There are three possible cases to consider. In each case, $E_1 = K_{E_1}\{\forall xA +^x A\}$ is an expansion tree with $Dp^+(E_1) \equiv K_1\{A\}$ and, by the inductive hypothesis, we have a derivation $\begin{array}{c}\forall xK_1\{A\}\\ \Vert\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\}\\ K_1\{\forall xA\}\end{array}$.

1. $E = (E_1 \star E_2)$, with $Dp^+(E) \equiv K_1\{A\} \star Dp^+(E_2)$. As e is minimal, it cannot point to any edge in E_2. Therefore $B :\equiv Dp^+(E_2)$ is free for x. Therefore we can construct the derivations:

$$\mathsf{r1}\downarrow\frac{\forall x(K_1\{A\} \vee B)}{\begin{array}{c}\forall xK_1\{A\}\\ \Vert\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\} \vee B\\ K_1\{\forall xA\}\end{array}} \quad \textit{and} \quad \mathsf{r2}\downarrow\frac{\forall x(K_1\{A\} \wedge B)}{\begin{array}{c}\forall xK_1\{A\}\\ \Vert\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\} \wedge B\\ K_1\{\forall xA\}\end{array}}$$

2. $E = \forall y(Sh(E_1)) +^y E_1$. As $Dp^+(E) \equiv Dp^+(E_1)$, we are already done.

3. $E = \exists yK_0\{A_0\} +^{t_1} E_1 \cdots +^{t_n} E_n$, with $Dp^+(E_i) \equiv B_i :\equiv [t_i \Rightarrow y](K_0\{A_0\})$ and in particular $B_1 \equiv K_1\{A\}$. Thus $Dp^+(E) \equiv \exists yB_0 \vee K_1\{A\} \vee B_2 \vee \ldots \vee B_n$. Again, e cannot point to any edge in any of the E'_i, so we can construct:

$$\mathsf{r1}\downarrow\frac{\forall x(\exists yB_0 \vee K_1\{A\} \vee B_2 \vee \ldots \vee B_n)}{\mathsf{r1}\downarrow\dfrac{\forall x(\exists yB_0 \vee K_1\{A\})}{\left(\exists yB_0 \vee \begin{array}{c}\forall xK_1\{A\}\\ \Vert\{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\}\\ K_1\{\forall xA\}\end{array}\right)} \vee (B_2 \vee \ldots \vee B_n)}$$

□

Lemma 4.45. *Let $E = E_0 + \bot(E_1,\ldots,E_n)$, with expansive deep formula $Dp^+(E) \equiv Dp^+(E_0) \vee Dp^+(E_1) \vee \ldots \vee Dp^+(E_n)$, $Dp^+(E_i) \equiv A_i$, and $A_k \equiv K\{B\}$ for some particular $0 \leq k \leq n$, be a correct weak expansion tree with cut, s.t. the $\forall$-edge labelled by x (which we will call e) is minimal w.r.t. $<_E$.*

Then there is a derivation:

$$\begin{array}{c} \forall x A_0 \vee A_1 \vee \ldots \vee K\{B\} \vee A_n \\ \| \{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\} \\ A_0 \vee A_1 \vee \ldots \vee K\{\forall x B\} \vee A_n \end{array}$$

Proof. Since x is minimal w.r.t. $<_E$, it is certainly minimal w.r.t $<_{E_k}$. Therefore, by Lemma 4.44, we can construct the derivation $\begin{array}{c} \forall x K\{B\} \\ \| \{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\} \\ K\{\forall x B\} \end{array}$. Therefore, we can also construct the derivation:

$$\begin{array}{c} \forall x(A_0 \vee A_1 \vee \ldots \vee A_k \vee \ldots \vee A_n) \\ \| \{\mathsf{r1}\downarrow\} \\ A_0 \vee A_1 \vee \ldots \vee \boxed{\begin{array}{c} \forall x K\{B\} \\ \| \{\mathsf{r1}\downarrow,\mathsf{r2}\downarrow\} \\ K\{\forall x B\} \end{array}} \vee \ldots \vee A_n \end{array}$$

□

We are now ready to define the map form expansion proofs to HNF proofs.

Definition 4.46. We define the map π_2^{Lo} : EPC $\rightarrow$ HNFC:

$$\pi_2^{Lo}(E) \equiv \begin{array}{c} \forall \vec{x} Dp^+(E) \\ \| \{\mathsf{h}\downarrow,\mathsf{r1}\downarrow,\mathsf{r2}\downarrow,qiU\} \\ Sh(E) \end{array}$$

- If E is just a leaf A, $\pi_2^{Lo}(E) \equiv A$.
- If $E = K_E\{B_1 \star_{E_1} C_1\}\ldots\{B_n \star_{E_n} C_n\}$, where E_i are all the $\star$-nodes s.t. the edges between $\star_{E_i}$ and B_i and between $\star_{E_i}$ and C_i are minimal, then we define $\pi_2^{Lo}(E) \equiv \pi_2^{Lo}(E')$, with $E' = K_E\{B_1 \star_{F_1} C_1\}\ldots\{B_n \star_{F_n} C_n\}$, which is a correct weak expansion tree with cut. Pictorially:

$$E = K_E \left\{ \begin{array}{ccc} B_1 & & C_1 \\ & \diagdown\ \diagup & \\ & \star & \end{array} \right\} \cdots \left\{ \begin{array}{ccc} B_n & & C_n \\ & \diagdown\ \diagup & \\ & \star & \end{array} \right\}$$

$$E' = K_E\{B_1 \star C_1\}\ldots\{B_n \star C_n\}$$

- Assume E has no minimal $\star$ edges. If

$$E = K_E\{\exists x_1 A_1 +^{t_1^1} E_1^1 \cdots +^{t_1^{m_1}} E_1^{m_1}\}\ldots\{\exists x_n A_n +^{t_n^1} E_n^1 \cdots +^{t_n^{m_n}} E_n^{m_n}\}$$

with

$$Dp^+(E) \equiv K\{\exists x A_1 \vee A_1^1 \vee \ldots \vee A_1^{m_1}\}\ldots\{\exists x A_n \vee A_n^1 \vee \ldots \vee A_n^{m_n}\}$$

and all edges $e_i^{j_i}$ minimal for $1 \leq i \leq n$ and $k_i < j_i \leq m_i$ (where $1 \leq k_i \leq m_i$), then

$$E' = K_E\{\exists x_1 A_1 +^{t_1^1} E_1^1 \cdots +^{t_1^{k_1}} E_1^{k_1}\} \ldots \{\exists x_n A_n +^{t_n^1} E_n^1 \cdots +^{t_n^{k_n}} E_n^{k_n}\}$$

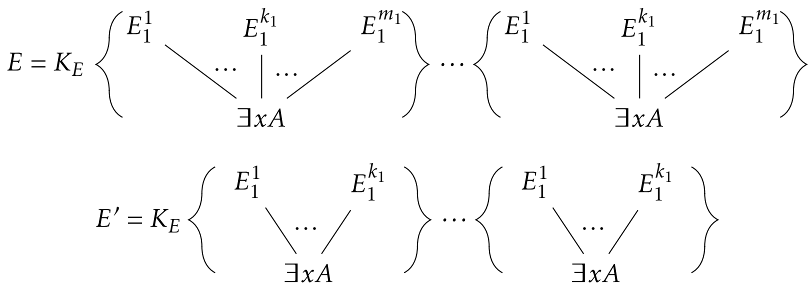
is a correct weak expansion tree with cut with

$$Dp^+(E) \equiv K\{\exists x A_1 \vee A_1^1 \vee \ldots \vee A_1^{k_1}\} \ldots \{\exists x A_n \vee A_n^1 \vee \ldots \vee A_n^{k_n}\}$$

and we can define:

$$\pi_2^{Lo}(E) \equiv \frac{K\left\{\begin{array}{c}\exists x A_1 \vee A_1^1 \vee \ldots \vee A_1^{m_1} \\ \| \{\mathsf{h}\downarrow\} \\ \exists x A_1 \vee A_1^1 \vee \ldots \vee A_1^{k_1}\end{array}\right\} \ldots \left\{\begin{array}{c}\exists x A_n \vee A_n^1 \vee \ldots \vee A_n^{m_n} \\ \| \{\mathsf{h}\downarrow\} \\ \exists x A_n \vee A_n^1 \vee \ldots \vee A_n^{k_n}\end{array}\right\}}{\pi_2^{Lo}(E')} =$$

Pictorially:



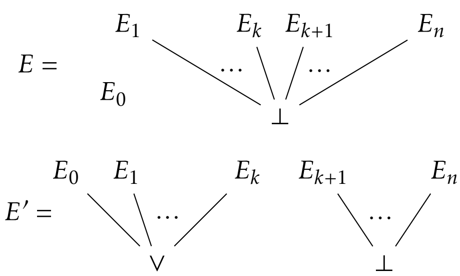
- Assume E does not have any minimal $\star$ nodes or $\exists$ edges. Let $E = E_0 + \perp(E_1, \ldots, E_k, \ldots, E_n)$ with the cut edges e_j for $1 \leq j \leq k$ all minimal. Then $E' = (E_0 \vee E_1 \vee \ldots \vee E_k) + \perp(E_{k+1}, \ldots, E_n)$ is a correct (weak) expansion tree with cut with

$$Sh(E') \equiv (Sh(E_0) \vee Sh(E_1) \vee \ldots \vee Sh(E_k)) \text{ and } Dp^+(E') \equiv Dp^+(E)$$

and we can define

$$\pi_2^{Lo}(E) \equiv \frac{\dfrac{\pi_2^{Lo}(E')}{A \vee \mathsf{qi}\uparrow \dfrac{A_1 \wedge \bar{A}_1}{\mathsf{f}} \vee \ldots \vee \mathsf{qi}\uparrow \dfrac{A_k \wedge \bar{A}_k}{\mathsf{f}}}}{A} =$$

Pictorially:



- Assume E is a weak expansion proof with cut with no minimal $\star$ or cut node, and no minimal $\exists$ edge. Then $E = K_E\{\forall xA +^x A\}$ for some minimal $\forall$ node, and by Lemma 4.45, $E' = K_E\{\forall xA\}$ is a correct weak expansion tree with cut and we can define:

$$\pi_2^{Lo}(E) \equiv \begin{array}{c} \forall x Dp^+(E) \\ \Vert \{\mathsf{r1}{\downarrow},\mathsf{r2}{\downarrow}\} \\ Dp^+(E') \\ \overset{=}{\cdots\cdots\cdots} \\ \pi_2^{Lo}(E') \end{array}$$

Pictorially:

$$E = K_E \left\{ \begin{array}{c} A \\ | \\ \forall xA \end{array} \right\} \qquad\qquad E' = K_E\{\forall xA\}$$

Theorem 4.47. *If E is an expansion proof with cut where $Sh(E) \equiv A$, then we can construct an* KSh2c *proof ϕ of A in HNFC, where $H_\phi(A) \equiv Dp(E)$.*

Proof. As $Dp(E)$ is a tautology, there is a proof $\begin{array}{c} \pi_2^{Up}(E) \Vert \mathsf{KS} \\ \forall\vec{x}Dp(E) \end{array}$ and clearly there is a proof $\begin{array}{c} Dp(E) \\ \Vert \{\exists\mathsf{w}{\downarrow}\} \\ Dp^+(E) \end{array}$. Thus, assuming we have some strategy for picking minimal $\forall$-nodes, we can define π_2 from expansion proofs to KSh2c proofs in HNF as:

$$\pi_2(E) \equiv \begin{array}{c} \pi_2^{Up}(E) \Vert \mathsf{KS} \\ \forall\vec{x}Dp(E) \\ \Vert \{\exists\mathsf{w}{\downarrow}\} \\ \forall\vec{x}Dp^+(E) \\ \pi_2^{Lo}(E) \Vert \{\mathsf{r1}{\downarrow},\mathsf{r2}{\downarrow},\mathsf{h}{\downarrow},\mathsf{qi}{\uparrow}\} \\ Sh(E) \end{array}$$

□

Observation 4.48. For all expansion proofs with cut E we have:

$$\pi_2^{Up}(E) \equiv Up(\pi_2(E)) \quad \text{and} \quad \pi_2^{Lo}(E) \equiv Lo(\pi_2(E)).$$

Remark 4.49. Although π_2 as defined here is a big improvement on the π_2 defined in [Ral18], there is still a small element of choice involved. If one thinks game semantically, π_2 is equivalent to constructing a proof by ∃loise playing every possible move on her turn (it is fairly obvious that it doesn't make any significant difference in which order she makes these moves), followed by ∀belard choosing one possible move on his. Clearly which move ∀belard chooses affects the proof that will be constructed. Still, what we might call "∃loise canonicity" is an advance on what is possible in the sequent calculus, unless one adds some extra syntax, such as focussing [CHM16].

We could make progress towards "∀belard canoncity" by replacing r1↓ and r2↓ with a general retract rule, such as in [Brü06a], but then we lose a certain amount of fine-grainedness in the proofs.

The translation for expansion proofs with closed cut is actually a lot more straightforward, since we can separate the cuts from h↓, r1↓ and r2↓.

Corollary 4.50. *If E is an expansion proof with closed cuts s.t. $Sh(E) \equiv A$, then we can construct a proof*

$$\pi_3(E) \equiv \begin{array}{c} \pi_3^{Up}(E) \Big\| \mathsf{KS} \\ \forall \vec{x} Dp(E) \\ \| \{\exists \mathsf{w}\downarrow\} \\ \forall \vec{x} Dp^+(E) \\ \pi_3^{Lo}(E) \Big\| \{\mathsf{r1}\downarrow, \mathsf{r2}\downarrow, \mathsf{h}\downarrow\} \\ A \vee B \\ \| \{\mathsf{qi}\uparrow\} \\ A \end{array}$$

where $B \equiv (\forall x A_1 \wedge \exists x A_1) \vee \ldots \vee (\forall x A_n \wedge \exists x A_n)$

Proof. Instead of translating the expansion proof with cut, we replace the $\perp$ node with a series of $\vee$ nodes, to give an expansion proof E' with $Sh(E') \equiv Sh(E) \vee (\forall x A_1 \wedge \exists x A_1) \vee \ldots \vee (\forall x A_n \wedge \exists x A_n)$ and $Dp(E') \equiv Dp(E)$. Then, we just take

$$\pi_3(E) \equiv \begin{array}{c} \pi_2(E') \Big\| \{\mathsf{h}\downarrow, \mathsf{r1}\downarrow, \mathsf{r2}\downarrow\} \\ \boxed{A \vee \mathsf{qi}\uparrow \dfrac{\forall x A_1 \wedge \exists x A_1}{\mathsf{f}} \vee \ldots \vee \mathsf{qi}\uparrow \dfrac{\forall x A_n \wedge \exists x A_n}{\mathsf{f}}} \end{array}$$

□

Of course, if the expansion proof is cut free, so is the deep inference proof.

Corollary 4.51. *Let E be a cut-free expansion proof with $Sh(E) \equiv A$. Then we can construct a proof ϕ_E in HNF of A.*

Proof. Clearly $\pi_2(E)$ is cut-free if E is. □

Having translations back and forth between expansion proofs and deep inference proofs gives us access to simple ways to prove certain properties. For example, we can show how to eliminate switches from the lower part of HNF proofs.

Proposition 4.52. *The switch rule is admissible for the lower part of an HNF proof, i.e. if there is a proof in HNF,*

$$\phi \equiv \begin{array}{c} Up(\phi) \Big\| \mathsf{KS} \\ H_\phi(A \wedge (B \vee C)) \\ \| \{\exists \mathsf{w}\downarrow\} \\ H_\phi^+(A \wedge (B \vee C)) \\ Lo(\phi) \| \{\mathsf{r1}\downarrow, \mathsf{r2}\downarrow, \mathsf{h}\downarrow, \mathsf{qi}\uparrow\} \\ K\{A \wedge (B \vee C)\} \end{array}$$

we can construct:

$$\phi' \equiv \begin{array}{c} Up(\phi') \,\Pi\, \mathsf{KS} \\ H_{\phi'}((A \wedge B) \vee C) \\ \Vert \{\exists \mathsf{w}\!\downarrow\} \\ H^{+}_{\phi'}((A \wedge B) \vee C) \\ Lo(\phi') \Vert \{\mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow, \mathsf{h}\!\downarrow, \mathsf{qi}\!\uparrow\} \\ K\{(A \wedge B) \vee C\} \end{array}$$

Proof. Take $E_\phi = \pi_1(\phi)$. $E_\phi = K_E\{E_A \wedge (E_B \vee E_C)\}$, with $Sh(E_A) \equiv A$, $Sh(E_B) \equiv C$ and $Sh(E_C) \equiv C$. Define $E' = K_E\{(E_A \wedge E_B) \vee E_C\}$. Clearly $Sh(E') \equiv K\{(A \wedge B) \vee C\}$. We need to check if E' is correct. Clearly, any dependency cycle in E' could easily be transformed into a cycle in E. We have from ϕ a proof $\begin{array}{c} \Pi\,\mathsf{KS} \\ K'\{A' \wedge (B' \vee C')\} \end{array}$ where $A' \equiv Dp(A)$, $B' \equiv Dp(B)$ and $C' \equiv Dp(C)$. Therefore we have $Up(\phi') \equiv$ $\begin{array}{c} \Pi\,\mathsf{KS} \\ K'\left\{ \mathsf{s}\dfrac{A' \wedge (B' \vee C')}{(A' \wedge B') \vee C'} \right\} \end{array}$, so E' is correct. Therefore we can construct $\phi' \equiv \pi_2(E')$.

□

Corollary 4.53. $\mathsf{i}\!\uparrow$ *is admissible for proofs in HNFC when in the lower part, i.e, if we have a proof*

$$\phi \equiv \begin{array}{c} \Pi\,\mathsf{KS} \\ H_{\phi}(A) \\ \Vert \{\exists \mathsf{w}\!\downarrow\} \\ H^{+}_{\phi}(A) \\ \Vert \{\mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow, \mathsf{h}\!\downarrow, \mathsf{qi}\!\uparrow, \mathsf{i}\!\uparrow\} \\ A \end{array}$$

we can construct a proof in HNFC

$$\phi' \equiv \begin{array}{c} \Pi\,\mathsf{KS} \\ H_{\phi'}(A) \\ \Vert \{\exists \mathsf{w}\!\downarrow\} \\ H^{+}_{\phi'}(A) \\ \Vert \{\mathsf{r1}\!\downarrow, \mathsf{r2}\!\downarrow, \mathsf{h}\!\downarrow, \mathsf{qi}\!\uparrow\} \\ A \end{array}$$

Proof. As seen previously $\mathsf{i}\!\uparrow$ is derivable for $\{\mathsf{ai}\!\uparrow, \mathsf{qi}\!\uparrow, \mathsf{s}\}$. We can simply push instances of $\mathsf{ai}\!\uparrow$ up through the lower part of the proof, and eliminate them in the upper part by propositional cut elimination.

By Proposition 4.52, we can eliminate any switches that are generated. □

4.4 Cut Elimination for Expansion Proofs

Unfortunately, providing a new cut elimination procedure for expansion proofs has proven to be beyond the scope of this thesis. However, there are a number of "off the shelf" procedures in the literature. McKinley's Herbrand nets are proof nets for the sequent calculus, and so Herbrand net cut reductions adhere closely to those in the sequent calculus [McK13]. The cut reductions for Heijltjes's proof forests diverge from the sequent calculus, borrowing more from game semantical techniques [Hei10]. However, a key ingredient for weak normalisation is a different correctness condition to standard expansion trees, and thus it is not clear that some of the techniques made possible by this adjusted correctness condition—such as the *pruning* of proof forests—would translate naturally into either sequent calculus or deep inference. The unpublished work of Aschieri et al. gives a much more syntactic cut elimination procedure for Miller-style expansion proofs [AHW18]. Unlike in McKinley and Heijltjes's papers, expansion trees are not limited to prenex formulae, although there is no *prima facie* reason why an extension to all first-order formulae would not be possible for these cut elimination procedures as well.

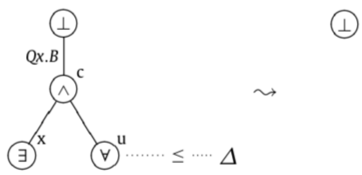
Theorem 4.54. *Let E be an expansion proof with cut. We can obtain a cut-free expansion proof E' with $Sh(E') = Sh(E)$.*

Proof. Using the techniques from [Hei10], [McK13] or [AHW18]. □

We will now look more closely at cut reduction for expansion proofs, and how it might be interpreted in HNFC proofs. We will follow the taxonomy of Heijltjes, who identifies four reduction steps: propositional, disposal, logical, and structural. We will look at each of them briefly, for now ignoring the variable conditions, and seeing what they would naively correspond to in a deep inference system.

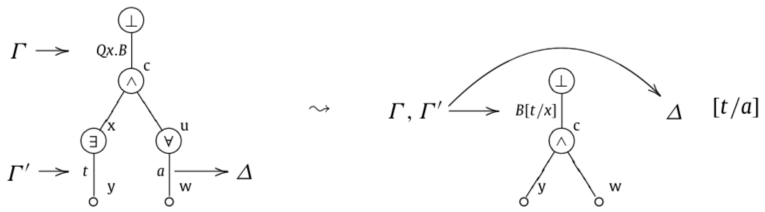
The propositional step corresponds to the elimination of quantifier-free cuts; we have already shown that we need not worry about these as they can be permuted into the upper part of a proof in HNFC and eliminated.

The disposal step can be imitated by a $\mathsf{w}\!\downarrow - \mathsf{qi}\!\uparrow$, followed by eliminating $\mathsf{w}\!\uparrow$ with a first-order version of the W reduction system:



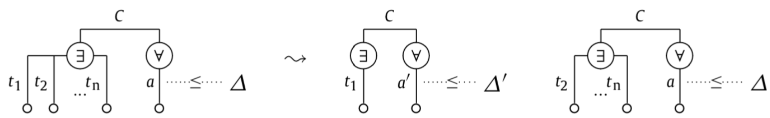
$$
\mathsf{qi}\!\uparrow \frac{\mathsf{w}\!\downarrow \dfrac{\mathsf{f}}{\exists x A} \wedge \forall x \bar{A}}{\mathsf{f}}
\quad \longrightarrow \quad
= \frac{\mathsf{f} \wedge \mathsf{w}\!\uparrow \dfrac{\forall x \bar{A}}{\mathsf{t}}}{\mathsf{f}}
$$

The logical step is equivalent to a $\mathsf{n}\!\downarrow - \mathsf{qi}\!\uparrow$ reduction, followed by eliminating the $\mathsf{n}\!\uparrow$ rule that is produced by, e.g. Lemma 3.22.



$$\mathsf{qi}\!\uparrow \dfrac{\mathsf{n}\!\downarrow\dfrac{[t\Rightarrow x]A}{\exists xA} \wedge \forall x\bar{A}}{\mathsf{f}} \quad\longrightarrow\quad \mathsf{i}\!\uparrow\dfrac{[t\Rightarrow x]A \wedge \mathsf{n}\!\downarrow\dfrac{\forall x\bar{A}}{[t\Rightarrow x]\bar{A}}}{\mathsf{f}}$$

The structural step is equivalent to a $\mathsf{qc}\!\downarrow - \mathsf{qi}\!\uparrow$ rewrite, followed by eliminating $\mathsf{qc}\!\uparrow$ with a first-order version of the C reduction system.

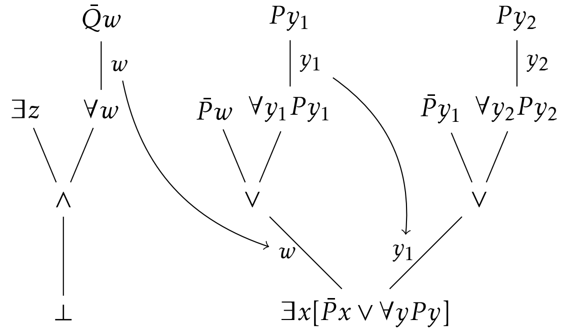


$$\mathsf{qi}\!\uparrow \dfrac{\mathsf{qc}\!\downarrow\dfrac{\exists xA \vee \exists xA}{\exists xA} \wedge \forall x\bar{A}}{\mathsf{f}} \quad\longrightarrow\quad \mathsf{2s}\dfrac{(\exists xA \vee \exists xA) \wedge \mathsf{qc}\!\uparrow\dfrac{\forall xA}{\forall xA \wedge \forall xA}}{\mathsf{qi}\!\uparrow\dfrac{\exists xA \wedge \forall x\bar{A}}{\mathsf{f}} \vee \mathsf{qi}\!\uparrow\dfrac{\exists xA \wedge \forall x\bar{A}}{\mathsf{f}}}$$

We will now discuss in more detail the disposal, logical and structural step, and the issues they present in the various versions of expansion trees with cut.

4.4.1 The disposal step

At first glance, this wouldn't seem to cause any issues. However, the version presented by Heijltjes for proof forests is in fact unsound for the expansion proofs as we present them. Consider the following expansion proof, created by using a witness from the universal quantifier opposite a weakened existential:



It is easily seen to be a correct expansion tree, and therefore we are able to construct a proof in HNFC corresponding to it:

$$\mathsf{r1}{\downarrow}\dfrac{\forall w\boxed{\mathsf{r1}{\downarrow}\dfrac{\forall y_1\forall y_2\boxed{\begin{array}{c}\Vert\mathsf{KS}\\(\mathsf{f}\wedge\bar{Q}w)\vee((Pw\vee\bar{P}y_1)\vee(Py_1\vee\bar{P}y_2))\end{array}}}{\left(\mathsf{\exists w}{\downarrow}\dfrac{\mathsf{f}}{\exists zQz}\wedge\bar{Q}w\right)\vee \mathsf{r1}{\downarrow}\dfrac{\forall y_1\boxed{\mathsf{r1}{\downarrow}\dfrac{\forall y_2((Pw\vee\bar{P}y_1)\vee(Py_1\vee\bar{P}y_2))}{(Pw\vee\bar{P}y_1)\vee\mathsf{h}{\downarrow}\dfrac{\forall y_2(Py_1\vee\bar{P}y_2)\vee\mathsf{\exists w}{\downarrow}\dfrac{\mathsf{f}}{\exists x\forall y(Px\vee\bar{P}y)}}{\exists x\forall y(Px\vee\bar{P}y)}}}}{\mathsf{h}{\downarrow}\dfrac{\forall y_1(Pw\vee\bar{P}y_1)\vee\exists x\forall y(Px\vee\bar{P}y)}{\exists x\forall y(Px\vee\bar{P}y)}}}}}{\mathsf{qi}{\uparrow}\dfrac{\mathsf{r2}{\downarrow}\dfrac{\forall w(\exists zQz\wedge\bar{Q}w)}{\exists zQz\wedge\forall w\bar{Q}w}}{\mathsf{f}}\vee\exists x\forall y(Px\vee\bar{P}y)}$$

However, if we perform the proof forestry disposal step, since all dependents of the tree opposite a weakened existential must be deleted, we are left with the following, clearly incorrect, expansion tree.

$$\exists x(\bar{P}x \vee \forall y Py)$$

So why is the disposal step not unsound for proof forests? The immediate reason is that the correctness condition is different: it is more tightly bound to the cut reduction procedure, and therefore both expansion trees, considered as proof forests, are incorrect. The moral reason *why* they are considered incorrect is because the class of expansion trees corresponding to HNFC proofs is slightly larger than the class of expansion trees corresponding to sequent calculus proofs. Whereas a (multiplicative) sequent calculus cut permanently splits the context in two, after which the two contexts cannot interact, in deep inference this is not the case. A sequent calculus proof corresponding to the above expansion tree would have to illegally weaken on an empty sequent to give the existential formula, whereas in deep inference this is not the case.

$$\text{Cut}\,\dfrac{\text{Weak}\,\dfrac{\vdash}{\vdash \exists zQz} \qquad \forall\text{R}\,\dfrac{\begin{array}{c}\mathrm{DF}\\ \vdash \bar{Q}w, \exists x(\bar{P}x \vee \forall yPy)\end{array}}{\vdash \forall w\bar{Q}w, \exists x(\bar{P}x \vee \forall yPy)}}{\vdash \exists x(\bar{P}x \vee \forall yPy)}$$

In [McK13], this problem is circumvented by simply disallowing weakening from the sequent system and insisting that each existential node have at least one branch (warning that "[w]eakening is notoriously difficult to handle well in proof nets"). In [AHW18] implicit weakenings on existential formulae correspond to existential nodes with one branch (with the label the quantified variable). The cut reduction involves renaming any dependent terms, and so is strictly different to the proof forestry disposal step.

Clearly, if we are to have a sound disposal step, we cannot use the disposal step as found in [Hei10]. Instead of deleting the dependent parts of eliminated tree, we must rename them. In HNFC, we can mimic this by permuting the $\exists\mathsf{w}{\downarrow}$ rule down and through the cut:

$$\exists\mathsf{w}{\downarrow}-\rho_K : \quad \rho\,\frac{K\left\{\exists\mathsf{w}{\downarrow}\dfrac{\mathsf{f}}{\exists xA}\right\}}{K'\{\exists xA\}} \quad\longrightarrow\quad \rho\,\frac{K\{\mathsf{f}\}}{K'\left\{\exists\mathsf{w}{\downarrow}\dfrac{\mathsf{f}}{\exists xA}\right\}}$$

$$\exists\mathsf{w}{\downarrow}-\mathsf{qi}{\uparrow} : \quad \mathsf{qi}{\uparrow}\,\frac{\mathsf{w}{\downarrow}\dfrac{\mathsf{f}}{\exists xA} \wedge \forall x\bar{A}}{\mathsf{f}} \quad\longrightarrow\quad =\,\frac{\mathsf{f} \wedge \mathsf{w}{\uparrow}\dfrac{\forall x\bar{A}}{\mathsf{t}}}{\mathsf{f}}$$

Instead of pushing the coweakening up as one, we reduce it to atomic form, pushing the atomic coweakenings up the proof using W (no new rewrites are needed), and remove the vacuous quantifiers. We can then replace all the universally quantified variables with fresh constants.

4.4.2 Logical Step

Of course, since $\mathsf{KSh2}$ does not contain $\mathsf{qc}{\downarrow}$ or $\mathsf{n}{\downarrow}$, but a single rule that combines both, $\mathsf{h}{\downarrow}$, there is not as clean a distinction between the logical and structural steps as there is in [Hei10]. In fact, [AHW18] and [McK13] make no distinction between them, treating the reductions generally. We will separate the two steps to focus on two issues that can be seen as particular to the logical step and the structural step: bridges and duplication, respectively.

First, we deal with bridges. A naive cut elimination strategy would involve permuting $\mathsf{h}{\downarrow}$ rules down until they hit a cut.

$$\mathsf{h}{\downarrow}-\mathsf{r1}{\downarrow}_1 : \quad \mathsf{r1}{\downarrow}\,\frac{\forall y\boxed{\mathsf{h}{\downarrow}\dfrac{\exists xA \vee [t \Rightarrow x]A}{\exists xA} \vee B}}{\forall y\exists xA \vee B} \quad\longrightarrow\quad \mathsf{r1}{\downarrow}\,\frac{\forall y((\exists xA \vee [t \Rightarrow x]A) \vee B)}{\forall y\boxed{\mathsf{h}{\downarrow}\dfrac{\exists xA \vee [t \Rightarrow x]A}{\exists xA}} \vee B}$$

$$\mathsf{h}{\downarrow}-\mathsf{r1}{\downarrow}_2 : \quad \mathsf{r1}{\downarrow}\,\frac{\forall y\boxed{\mathsf{h}{\downarrow}\dfrac{\exists xA \vee [t \Rightarrow x]A}{\exists xA} \vee B}}{\exists xA \vee \forall yB} \quad\longrightarrow\quad \mathsf{r1}{\downarrow}\,\frac{\forall y((\exists xA \vee [t \Rightarrow x]A) \vee B)}{\mathsf{h}{\downarrow}\dfrac{\exists xA \vee [t \Rightarrow x]A}{\exists xA} \vee \forall yB}$$

$$\mathsf{h}{\downarrow}-\mathsf{r2}{\downarrow}_1 : \quad \mathsf{r2}{\downarrow}\,\frac{\forall y\boxed{\mathsf{h}{\downarrow}\dfrac{\exists xA \vee [t \Rightarrow x]A}{\exists xA} \wedge B}}{\forall y\exists xA \wedge B} \quad\longrightarrow\quad \mathsf{r2}{\downarrow}\,\frac{\forall y((\exists xA \vee [t \Rightarrow x]A) \wedge B)}{\forall y\boxed{\mathsf{h}{\downarrow}\dfrac{\exists xA \vee [t \Rightarrow x]A}{\exists xA}} \wedge B}$$

$$\mathsf{h}{\downarrow}-\mathsf{r2}{\downarrow}_2 : \quad \mathsf{r2}{\downarrow}\,\frac{\forall y\boxed{\mathsf{h}{\downarrow}\dfrac{\exists xA \vee [t \Rightarrow x]A}{\exists xA} \wedge B}}{\exists xA \wedge \forall yB} \quad\longrightarrow\quad \mathsf{r2}{\downarrow}\,\frac{\forall y((\exists xA \vee [t \Rightarrow x]A) \wedge B)}{\mathsf{h}{\downarrow}\dfrac{\exists xA \vee [t \Rightarrow x]A}{\exists xA} \wedge \forall yB}$$

Note, however, that for $\mathsf{h}\!\downarrow - \mathsf{r1}\!\downarrow_2$ and $\mathsf{h}\!\downarrow - \mathsf{r2}\!\downarrow_2$, if t contains y, these rewrites will be invalid, thus stalling cut elimination. If the expansion proof is generated from an HNFC proofs by π_2, this configuration will not be common, since all possible cut, existential and binary connective "moves" will be "played" simultaneously before any universal moves. However, it is a possible configuration that can arise, in particular if we have an existential node that points to a universal node the other side of cut to it. This is the problem identified as *bridges* in [Hei10].

Example 4.55. Below is part of a HNFC and it's corresponding expansion tree that resists normalisation due to a bridge.

$$\mathsf{r2}\!\downarrow\dfrac{\forall x\,\boxed{\mathsf{h}\!\downarrow\dfrac{\exists y Py \vee Pt(x)}{\exists y Py} \wedge \bar{P}x}}{\mathsf{qi}\!\uparrow\dfrac{\exists y Py \wedge \forall x \bar{P}x}{\mathsf{f}}}
\qquad
\begin{tikzcd}
Pt(x) \arrow[d, dash] & t(x) \longleftarrow x & \bar{P}x \arrow[d, dash] \\
\exists y \arrow[dr, dash] & & \forall x \arrow[dl, dash] \\
& \wedge \arrow[d, dash] & \\
& \bot &
\end{tikzcd}$$

Note that if we set $t(x) = x$, then this is a dualised version of the proof of the drinker's formula.

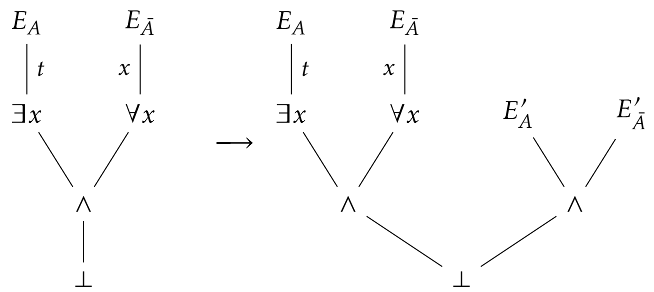
The solution to bridges in [Hei10] is for the correctness condition to essentially ignore any tree with a bridge. Thus they can be "pruned" safely. [McK13] avoids the problem altogether, by adapting the cut elimination procedure to make sure that bridges can never be formed in Herbrand nets. Since bridges are not possible in expansion proofs translated from sequent calculus proofs, due to the eigenvariable condition, the class of Herbrand nets considered can exclude the possibility of bridges. In [AHW18], bridges are broken by the renaming policy in their cut reduction step. It is not clear which of these strategies, if any, can make a suitable transition into $\mathsf{KSh2c}$ proofs, although, since pruning relies on a different correctness condition to the expansion tree norm, it is unlikely that it could be translated with ease.

4.4.3 The Structural Step

Once the problem of bridges is solved, and we are able to permute $\mathsf{h}\!\downarrow$ rules to their respective cuts, we can permute the $\mathsf{h}\!\downarrow$ rule through the cut.

$$\mathsf{qi}\!\uparrow\dfrac{\mathsf{h}\!\downarrow\dfrac{\exists x A \vee [t \Rightarrow x]A}{\exists x A} \wedge \forall x \bar{A}}{\mathsf{f}}
\quad\longrightarrow\quad
\mathsf{2s}\dfrac{(\exists x A \vee [t \Rightarrow x]A) \wedge \mathsf{h}\!\uparrow\dfrac{\forall x \bar{A}}{\forall x \bar{A} \wedge [t \Rightarrow x]\bar{A}}}{\mathsf{qi}\!\uparrow\dfrac{\exists x A \wedge \forall x \bar{A}}{\mathsf{f}} \vee \mathsf{i}\!\uparrow\dfrac{[t \Rightarrow x]A \wedge [t \Rightarrow x]\bar{A}}{\mathsf{f}}}$$

We would then need to permute the $\mathsf{h}\!\uparrow$ rule up through the proof, duplicating the proof above the universal quantifier. We can sketch the operation on expansion trees that this would mirror, again ignoring dependencies for now.



We can get a better idea of how the operation might work by looking more closely at h$\uparrow$ elimination. The inference rule can be eliminated from HNF proofs, but at the cost of leaving contractions behind.

Lemma 4.56. h$\uparrow$ *is admissible for the lower part of proofs in HNF, leaving behind* c$\downarrow$ *instances.*

Proof. We can permute instances of h$\uparrow$ up a proof in the following way

$$\mathsf{r1}{\downarrow} - \mathsf{h}{\uparrow} : \quad \mathsf{r1}{\downarrow}\dfrac{\forall x(A\vee B)}{\mathsf{h}{\uparrow}\dfrac{\forall xA}{\forall xA\wedge[t\Rightarrow x]A}\vee B} \quad\longrightarrow\quad \mathsf{h}{\uparrow}\dfrac{\forall x(A\vee B)}{\mathsf{2s}\dfrac{\mathsf{r1}{\downarrow}\dfrac{\forall x(A\vee B)}{\forall xA\vee B}\wedge([t\Rightarrow x]A\vee B)}{(\forall xA\wedge[t\Rightarrow x]A)\vee\mathsf{c}{\downarrow}\dfrac{B\vee B}{B}}}$$

$$\mathsf{r2}{\downarrow} - \mathsf{h}{\uparrow} : \quad \mathsf{r2}{\downarrow}\dfrac{\forall x(A\wedge B)}{\mathsf{h}{\uparrow}\dfrac{\forall xA}{\forall xA\wedge[t\Rightarrow x]A}\wedge B} \quad\longrightarrow\quad \mathsf{h}{\uparrow}\dfrac{\forall x(A\wedge B)}{=\dfrac{\mathsf{r1}{\downarrow}\dfrac{\forall x(A\wedge B)}{\forall xA\wedge B}\wedge([t\Rightarrow x]A\wedge B)}{(\forall xA\wedge[t\Rightarrow x]A)\wedge\mathsf{w}{\uparrow}\dfrac{B\wedge B}{B}}}$$

$$\mathsf{h}{\downarrow} - \mathsf{h}{\uparrow}_1 : \quad \mathsf{h}{\uparrow}\dfrac{\forall yK\left\{\mathsf{h}{\downarrow}\dfrac{\exists xA\vee[t_1\Rightarrow x]A}{\exists xA}\right\}}{\forall yK\{\exists xA\}\wedge[t_2\Rightarrow y]K\{\exists xA\}} \quad\longrightarrow$$

$$\mathsf{qc}{\uparrow}\dfrac{\forall yK\{\exists xA\vee[t_1\Rightarrow x]A\}}{\forall yK\left\{\mathsf{h}{\downarrow}\dfrac{\exists xA\vee[t_1\Rightarrow x]A}{\exists xA}\right\}\wedge\forall yK\left\{\mathsf{h}{\downarrow}\dfrac{\exists xA\vee[t_1\Rightarrow x]A}{\exists xA}\right\}}$$

$$\mathsf{h}{\downarrow} - \mathsf{h}{\uparrow}_2 : \quad \mathsf{h}{\downarrow}\dfrac{\exists xK\{\forall yA\}\vee[t_1\Rightarrow x]K\{\forall yA\}}{\exists xK\left\{\mathsf{h}{\uparrow}\dfrac{\forall yA}{\forall yA\wedge[t_2\Rightarrow y]A}\right\}} \quad\longrightarrow$$

$$\mathsf{h}{\downarrow}\dfrac{\exists xK\left\{\mathsf{h}{\uparrow}\dfrac{\forall yA}{\forall yA\wedge[t_2\Rightarrow y]A}\right\}\vee[t_1\Rightarrow x]K\left\{\mathsf{h}{\uparrow}\dfrac{\forall yA}{\forall yA\wedge[t_2\Rightarrow y]A}\right\}}{\exists xK\{\forall yA\wedge[t_2\Rightarrow y]A\}}$$

Checking the variable conditions in $\mathsf{h}{\downarrow}\text{−}\mathsf{h}{\uparrow}_1$ and $\mathsf{h}{\downarrow}\text{−}\mathsf{h}{\uparrow}_2$ is essentially the same as in the proof for Proposition 3.22.

Clearly, these rewrites leave $\mathsf{s}, \mathsf{c}{\downarrow}$ and $\mathsf{w}{\uparrow}$ instances in the proof. We have shown that switch and coweakening are admissible, leaving $\mathsf{c}{\downarrow}$. □

At this stage it is not exactly clear how best to deal with the contractions—the obvious direction for a solution is for them to be absorbed into the *hD* rules somehow. Clearly, eliminating them will involve duplicating part of the expansion proof. But, as McKinley notes, there are choices to make about how much of the expansion proof/proof forest should be duplicated by the reduction step. The modified correctness condition in [Hei10] allows for a smaller amount of copying than in [McK13] or [AHW18], but the fine-grainedness of using $\mathsf{r1}{\downarrow}$ and $\mathsf{r2}{\downarrow}$ instead of eigenvariables could also give us access to a more compact cut reduction relation. On the other hand, the fact that HNF proofs include a prenexification, forcing a total ordering on the quantifiers, might mean that superfluous information is duplicated.

4.5 Cut Elimination for SKSq

Making use of a cut elimination proof for expansion proofs, we can now prove the main theorem of the thesis.

Theorem 4.57 (Cut elimination for SKSq)**.** *The first-order proof systems* SKSq *and* KSq *are weakly equivalent, i.e. if there is a proof with cut* $\begin{array}{c}\phi \,\Vert\, \mathsf{SKSq}\\ A\end{array}$ *then there is a cut-free proof* $\begin{array}{c}\psi \,\Vert\, \mathsf{KSq}\\ A\end{array}$.

Proof. Let $\begin{array}{c}\phi \,\Vert\, \mathsf{SKSq}\\ A\end{array}$. By Proposition 3.29 we can reduce all the up-rules to $\mathsf{ai}\uparrow$ and $\mathsf{qi}\uparrow$, pushing all instances of $\mathsf{qi}\uparrow$ to the bottom of the proof.

By Theorem 4.5, taking into account Remark 4.8, we can construct a Herbrand proof ϕ_2 of $A \wedge B$ where $B \equiv \left(\forall x_1 B_1 \wedge \exists x_1 \bar{B}_1\right) \vee \ldots \vee \left(\forall x_n B_n \wedge \exists x_n \bar{B}_n\right)$.

By Theorem 1.55, we can eliminate $\mathsf{ai}\uparrow$ from the upper part of the Herbrand Proof to form ϕ_3. By Proposition 4.29, we can construct a proof ϕ_4 of $A \wedge B$ in HNF.

By Theorem 4.38, we can construct an expansion proof with cut EC_{ϕ_4}. By Theorem 4.54, we can eliminate the cuts from EC_{ϕ_4} to give E_{ϕ_5}, and then translate back into a proof in HNF of A, ϕ_5 by Theorem 4.51. By Proposition 4.29, we can translate this back into a $\mathsf{KSh1}$ proof ϕ_6, and finally back into KSq with Proposition 4.9.

$$
\begin{array}{ccccccc}
\begin{array}{c}\phi \,\Vert\, \mathsf{SKSq}\\ A\end{array} & \xrightarrow{\text{Prop 3.29}} & \begin{array}{c}\phi_1 \,\Vert\, \mathsf{KSq}\cup\{\mathsf{ai}\uparrow\}\\ A \wedge B \\ \Vert\, \{\mathsf{qi}\uparrow\} \\ A\end{array} & \xrightarrow{\text{Thm 1.55}} & \begin{array}{c}\phi_2 \,\Vert\, \mathsf{KSh1}\cup\{\mathsf{ai}\uparrow\}\\ A \wedge B \\ \Vert\, \{\mathsf{qi}\uparrow\} \\ A\end{array} & \xrightarrow{\text{Thm 4.5}} & \begin{array}{c}\phi_3 \,\Vert\, \mathsf{KSh1}\\ A \wedge B \\ \Vert\, \{\mathsf{qi}\uparrow\} \\ A\end{array} \\
 & \xrightarrow{\text{Prop 4.29}} & \begin{array}{c}\phi_4 \,\Vert\, \mathsf{KSh2}\\ A \wedge B \\ \Vert\, \{\mathsf{qi}\uparrow\} \\ A\end{array} & \xrightarrow{\text{Thm 4.38}} & \begin{array}{c}EC_{\phi_4} \,\Vert\, EPC\\ A\end{array} & \xrightarrow{\text{Thm 4.54}} & \begin{array}{c}E_{\phi_5} \,\Vert\, EP\\ A\end{array} \\
 & \xrightarrow{\text{Thm 4.47}} & \begin{array}{c}\phi_5 \,\Vert\, \mathsf{KSh2}\\ A\end{array} & \xrightarrow{\text{Prop 4.29}} & \begin{array}{c}\phi_6 \,\Vert\, \mathsf{KSh1}\\ A\end{array} & \xrightarrow{\text{Prop 4.9}} & \begin{array}{c}\psi \,\Vert\, \mathsf{KSq}\\ A\end{array}
\end{array}
$$

□

Conclusion

In this thesis, we have stated and proven three main theorems, each using a key proof-theoretic technology. In the first chapter, we prove propositional cut elimination by the experiments method, with the atomic flow providing an important framework. In the second chapter, decomposition for propositional proofs is reduced to cycle removal, which is performed with the use of a new rule: the merge contraction. Finally, in the fourth chapter, we show first-order cut elimination, with the key step being translations to and from expansion proofs. Of these three main strands, the last stands out in being the only one to substantially use logical technology developed outside the deep inference community. The first two results are, on the other hand, rather more parochial. Thus, we will suggest a few points of comparison with the wider logical literature, focussing in particular on the propositional material in the thesis, which has less direct contact with this literature. These comparisons will hopefully point towards future research directions, to the benefit of proof theory and logic more generally both within and without the deep inference community.

The Experiments Method and its Relations

A significant justification for the move to deep inference from Gentzen-style proof systems is the greater freedom in composing proofs and derivations. Using these freedoms, we are able to achieve a certain confluence in proof semantics for propositional logic, with methods such as the experiments. Thus, it will be useful to compare this method to techniques and innovations in more established proof systems that also aim at confluence of normalisation, or other means of achieving canonicity.

The Mix Rule

When eliminating cut from sequent calculus proofs, the standard technique for permuting the cut rule past a weakening on each cut formula is to delete one side of the proof and the cut with it. Since we have two choices, we have a clear counterexample to confluence of the cut elimination procedure.

$$\text{cut}\,\dfrac{\text{weak}\,\dfrac{\overset{\Pi_1}{\Gamma}}{\Gamma, A} \qquad \text{weak}\,\dfrac{\overset{\Pi_2}{\Delta}}{\bar{A}, \Delta}}{\Gamma, \Delta} \quad \Longrightarrow \quad \text{weak}\,\dfrac{\overset{\Pi_1}{\Gamma}}{\Gamma, \Delta} \quad / \quad \text{weak}\,\dfrac{\overset{\Pi_2}{\Delta}}{\Gamma, \Delta}$$

This situation, known as the Lafont counterexample [GTL89], is one where the formalism seems to force an arbitrary choice during normalisation, a problem that does not arise if we move to deep inference [Gug02]. Therefore we have access to cut elimination procedures such as the experiments method, which offer a form of canonicity unavailable to simple sequent systems.

Since, for the sequent calculus, the implicit connective between two proofs is always conjunction, there is seemingly no way to encode a normalisation method like the experiments, which relies on a disjunction on proofs. However, the addition of the *mix* rule goes someway towards solving this problem [Gen64; Pla01].

$$\text{Mix}\,\dfrac{\vdash \Gamma \quad \vdash \Delta}{\vdash \Gamma, \Delta}$$

The mix rule acts as a sort of meta level weakening, converting the implicit conjunction between two sequent calculus proofs into a disjunction at the level of the sequent. Thus we can envisage a version of the experiments method in the sequent calculus using the mix rule:

$$\text{Con}^{n-1}\,\dfrac{\text{Mix}^{n-1}\,\dfrac{\overset{\Delta_1}{\vdash \Gamma_1, B} \qquad \overset{\Delta_2}{\vdash \Gamma_2, B} \qquad \dots \qquad \overset{\Delta_n}{\vdash \Gamma_n, B}}{\vdash \Gamma_1, \Gamma_2, \dots, \Gamma_n, B, \dots, B}}{\vdash \Gamma_1, \Gamma_2, \dots, \Gamma_n, B}$$

Some important differences remain, though. First, although the mix rule can allow disjunctive rather than conjunctive connection of proofs, the structure of a sequent-calculus proof remains tree-like rather than fully two-dimensional. Thus, the truth-table tautology at the top of an experiments proof cannot be used in the same way, and this "experiments" method will not suffice as a cut elimination method as it currently is. Second, the mix rule does not change the implicit meta-connective between sequent calculus proofs, it only acts as a weakening of that connective from a conjunction to a disjunction at the point at which those proofs converge. Each proof joined together needs a valid premise, unlike when two deep inference proofs are disjunctively joined.

Quine's Method

A more further afield comparison to the experiments method can be made to Quine's method for minimising Boolean functions (later extended by McCluskey) [Qui52; Qui55; MJ56]. The problem of finding the simplest Boolean function equivalent to any given function has obvious analogies with the search for canonical proofs. The key to Quine's method is finding the *prime implicants* of a formula ϕ, that is the formulae that imply ϕ while subsuming no shorter formulae that also imply ϕ. Finding the prime implicants of a formula (in disjunctive normal form) is carried out by alternating two steps: dropping any superfluous clauses; and finding and adding the *consensus* of two clauses, constructed by removing dual atoms from each clause and then joining them in conjunction.

Since Quine's method has a different goal to the experiments, a direct comparison is not reasonable. However, there are interesting superficial similarities. On a general level, both methods are searching for a canonical object: Boolean functions and cut-free proofs, respectively. On a more technical level, finding the consensus of two clauses involves creating a new clause which has an atom deleted from one clause and its dual deleted from another, just as we create two experiments by deleting alternately an atom and its dual from a proof, in a certain way.

Graphical proof systems

Next, we will discuss two interesting graphical proof systems developed in recent years: Yves Lafont's *interaction nets* and Dominic Hughes' *combinatorial proofs*. Each of these formalisms satisfy the Cook-Reckhow criteria for a formal proof system [CR79], that the proof system must be sound, complete and be checkable in polynomial time. Our interest in graphical proof systems stems from their suitability as "bureaucracy-free" canonical proof objects [Gir89; BL05]. These proof systems are interesting in their own right, of course, but also because of the possibility of inter-translatability with more syntactically burdened proof systems, with such translations often giving clear indication of what proofs it is reasonable to identify semantically.

Interaction nets

We will first introduce a comparison that is already recognised and studied: between atomic flows and interaction nets. Just as atomic flows are a graphical formalism that arise from classical deep inference proof theory, interaction nets were developed out of the proof theory of linear logic, specifically their proof nets [Laf89]. Much like the atomic flow, rewrites govern the behaviour of interaction nets, with many of the rewrites—duplication, for example— essentially the same as those for the atomic flow. However, an important difference should be stressed: whereas interaction nets form a local and strongly confluent model of computation, the atomic flow itself is not sufficient for computation, since they do not form a proof system according to the Cook-Reckhow criteria described above. Instead, the atomic flow gives us access to certain aspects of SKS proof theory that would not be feasible without it.

Additionally, the presentation of merge contractions as context contractions suggests a comparison with context sharing (or partial sharing) in sharing graphs for optimal reduction. Sharing graphs, a particular subclass of interaction nets, have been shown to implement optimal reduction in the lambda calculus, where both one value can be shared in multiple contexts, and multiple values can be shared in a single shared context [Lam90; GAL92]. Clearly, there is some superficial similarity to the context contractions (shown to be equivalent to merge contractions) here, where we could choose to distinguish between context contractions with non-trivial contexts (i.e. at least one hole) and regular contractions, mirroring the distinction between context sharing and value sharing. Thus, it might be interesting to investigate Curry-Howard correspondences from proof systems with merge contractions, building on the work already carried out on the atomic lambda calculus [GHP13].

A more formalised link between deep inference and interaction nets is provided by Stéphane Gimenez and George Moser: a Curry-Howard correspondence between a linear logic-based system and a typed variant of sharing graphs, [GM13]. Furthermore, preliminary work has been carried out investigating the possibility of *atomic* graphs, extending the Curry-Howard correspondence established between deep inference intuitionistic logic and the atomic lambda calculus to sharing graphs [GHP13; SS17].

Combinatorial Proofs

Finally, we discuss a second graphical proof system: Dominic Hughes's *combinatorial proofs* [Hug06]. Here, proof objects are almost purely algebraic: a proof of ϕ consists of a graph homomorphism $h : C \to G(\phi)$, where $G(\phi)$ is a graph associated with ϕ and C a coloured graph. Crucially for this comparison, $G(\phi)$ does not respect the tree structure of ϕ but represents the logical relations between atoms in a formula, similar to the relation webs of Guglielmi's BV [Gug07]. In this way, combinatorial proofs can be thought of as a form of "deep" inference, although it is perhaps more accurate to describe it as a formalism that transcends the "deep/shallow" binary.

Recently, work towards first-order combinatorial proofs has been published by Hughes [Hug18], as well as work by Straßburger extending the notion to that of *combinatorial flows* [Str17a]. Combinatorial proofs have also been investigated as canonical proofs across a range of propositional proof systems, including sequent calculus, analytic tableaux, and resolution [AS18]—if we are to suggest certain deep inference proof systems or classes of deep inference proofs as canonical, combinatorial proofs offer us a benchmark.

Bibliography

[AB16] Juan P Aguilera and Matthias Baaz. "Unsound Inferences Make Proofs Shorter". In: *arXiv preprint arXiv:1608.07703* (2016).

[Ack40] Wilhelm Ackermann. "Zur widerspruchsfreiheit der zahlentheorie". In: *Mathematische Annalen* 117.1 (1940), pp. 162–194. ISSN: 0025-5831.

[AHW18] Federico Aschieri, Stefan Hetzl, and Daniel Weller. "Expansion Trees with Cut". Unpublished Work. arXiv preprint. 2018. URL: https://arxiv.org/abs/1802.08076.

[Alc+17] Aurore Alcolei et al. "The true concurrency of Herbrand's theorem". Unpublished Work. Submitted. 2017.

[AS18] Matteo Acclavio and Lutz Straßburger. "From syntactic proofs to combinatorial proofs". In: *International Joint Conference on Automated Reasoning*. Springer, 2018, pp. 481–497.

[AT16] Andrea Aler Tubella. "A study of normalisation through subatomic logic". Original Version. Thesis. University of Bath, 2016. URL: https://aalertubelladotcom.files.wordpress.com/2017/02/thesis.pdf.

[AT17] Andrea Aler Tubella. "A study of normalisation through subatomic logic". Revised Version. Thesis. University of Bath, 2017. URL: https://aalertubelladotcom.files.wordpress.com/2017/02/thesis.pdf.

[ATG18] Andrea Aler Tubella and Alessio Guglielmi. "Subatomic Proof Systems: Splittable Systems". In: *ACM Trans. Comput. Logic* 19.1 (2018), pp. 1–33. ISSN: 1529-3785. DOI: 10.1145/3173544.

[ATGR17] Andrea Aler Tubella, Alessio Guglielmi, and Benjamin Ralph. "Removing Cycles from Proofs". In: *26th EACSL Annual Conference on Computer Science Logic (CSL 2017)*. Ed. by Valentin Goranko and Mads Dam. Vol. 82. Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017, 9:1–9:17. ISBN: 978-3-95977-045-3. DOI: 10.4230/LIPIcs.CSL.2017.9. URL: http://cs.bath.ac.uk/ag/p/RCP.pdf.

[ATGR18] Andrea Aler Tubella, Alessio Guglielmi, and Benjamin Ralph. "Removing Cycles with Merge Contractions". Unpublished Work. Unpublished Draft. 2018.

[Avr87] Arnon Avron. "A constructive analysis of RM". In: *The Journal of symbolic logic* 52.4 (1987), pp. 939–951. ISSN: 0022-4812.

[Avr91] Arnon Avron. "Hypersequents, logical consequence and intermediate logics for concurrency". In: *Annals of Mathematics and Artificial Intelligence* 4.3-4 (1991), pp. 225–248. ISSN: 1012-2443.

[Bel82] Nuel D Belnap. "Display logic". In: *Journal of philosophical logic* 11.4 (1982), pp. 375–417. ISSN: 0022-3611.

[BG09] Paola Bruscoli and Alessio Guglielmi. "On the Proof Complexity of Deep Inference". In: *ACM Transactions on Computational Logic (TOCL)* 10.2 (2009), p. 14. ISSN: 1529-3785. DOI: 10.1145/1462179.1462186. URL: <GotoISI>://WOS:000266012100007.

[BG16] Paola Bruscoli and Alessio Guglielmi. *On Analyticity in Deep Inference*. Web Page. 2016. URL: http://cs.bath.ac.uk/ag/p/ADI.pdf.

[BL05] Kai Brünnler and Stéphane Lengrand. "On two forms of bureaucracy in derivations". In: (2005).

[BL94] Matthias Baaz and Alexander Leitsch. "On skolemization and proof complexity". In: *Fundamenta Informaticae* 20.4 (1994), pp. 353–379. ISSN: 0169-2968.

[BLL18] Matthias Baaz, Alexander Leitsch, and Anela Lolic. "A Sequent-Calculus Based Formulation of the Extended First Epsilon Theorem". In: *International Symposium on Logical Foundations of Computer Science*. Springer, 2018, pp. 55–71.

[BM08] Kai Brünnler and Richard McKinley. "An algorithmic interpretation of a deep inference system". In: *International Conference on Logic for Programming Artificial Intelligence and Reasoning*. Springer, 2008, pp. 482–496.

[Bru+16] Paola Bruscoli et al. "Quasipolynomial Normalisation in Deep Inference Via Atomic Flows and Threshold Formulae". In: *Logical Methods in Computer Science* 12.1 (2016), 5:1–30. ISSN: 1860-5974. DOI: Artn510.2168/Lmcs-12(1:5)2016. URL: http://cs.bath.ac.uk/ag/p/QuasiPolNormDI.pdf.

[Brü03a] Kai Brünnler. *Deep inference and symmetry in classical proofs*. Logos Verlag, 2003. ISBN: 3832504486.

[Brü03b] Kai Brünnler. "Two restrictions on contraction". In: *Logic Journal of IGPL* 11.5 (2003), pp. 525–529. ISSN: 1367-0751. URL: http://cs.bath.ac.uk/ag/kai/RestContr.pdf.

[Brü06a] Kai Brünnler. "Cut elimination inside a deep inference system for classical predicate logic". In: *Studia Logica* 82.1 (2006), pp. 51–71. ISSN: 0039-3215. URL: http://cs.bath.ac.uk/ag/kai/q.pdf.

[Brü06b] Kai Brünnler. "Locality for classical logic". In: *Notre Dame Journal of Formal Logic* 47.4 (2006), pp. 557–580. ISSN: 0029-4527.

[Brü09] Kai Brünnler. "Deep sequent systems for modal logic". In: *Archive for Mathematical Logic* 48.6 (2009), pp. 551–577. ISSN: 0933-5846.

[BT01] Kai Brünnler and Alwen Tiu. "A local system for classical logic". In: *Logic for Programming, Artificial Intelligence, and Reasoning*. Vol. 2250. Springer, 2001, pp. 347–361. ISBN: 3540429573. DOI: 10.1007/3-540-45653-8_24.

[Bus91] Samuel R Buss. "The undecidability of k-provability". In: *Annals of Pure and Applied Logic* 53.1 (1991), pp. 75–102. ISSN: 0168-0072. DOI: 10.1016/0168-0072(91)90059-U.

[Bus95] Samuel R Buss. "On Herbrand's theorem". In: *Logic and Computational Complexity*. Springer, 1995, pp. 195–209. ISBN: 3540601783.

[Car02] Alessandra Carbone. "The cost of a cycle is a square". In: *The Journal of Symbolic Logic* 67.1 (2002), pp. 35–60. ISSN: 1943-5886. DOI: 10.2178/jsl/1190150028. URL: <GotoISI>://WOS:000174333800003.

[CGT08] Agata Ciabattoni, Nikolaos Galatos, and Kazushige Terui. "From axioms to analytic rules in nonclassical logics". In: *23rd Annual IEEE Symposium on Logic in Computer Science*. IEEE, 2008, pp. 229–240. ISBN: 0769531830.

[CHM16] Kaustuv Chaudhuri, Stefan Hetzl, and Dale Miller. "A multi-focused proof system isomorphic to expansion proofs". In: *Journal of Logic and Computation* 26.2 (2016), pp. 577–603. ISSN: 0955-792X. DOI: 10.1093/logcom/exu030. URL: http://dx.doi.org/10.1093/logcom/exu030.

[Coq95] Thierry Coquand. "A Semantics of Evidence for Classical Arithmetic". In: *Journal of Symbolic Logic* 60.1 (1995). Qt521 Times Cited:55 Cited References Count:15, pp. 325–337. ISSN: 0022-4812. DOI: Doi 10.2307/2275524. URL: <GotoISI>://WOS:A1995QT52100018.

[CR79] Stephen A Cook and Robert A Reckhow. "The relative efficiency of propositional proof systems". In: *The Journal of Symbolic Logic* 44.01 (1979), pp. 36–50. ISSN: 1943-5886.

[CRW14] Agata Ciabattoni, Revantha Ramanayake, and Heinrich Wansing. "Hypersequent and display calculi–a unified perspective". In: *Studia Logica* 102.6 (2014), pp. 1245–1294. ISSN: 0039-3215.

[Das14a] Anupam Das. "On the pigeonhole and related principles in deep inference and monotone systems". In: *Joint Meeting of the 23rd EACSL Annual Conference on Computer Science Logic (CSL) and the 29th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. 2014.

[Das14b] Anupam Das. "The complexity of propositional proofs in deep inference". Thesis. University of Bath, 2014.

[Das15] Anupam Das. "On the Relative Proof Complexity of Deep Inference Via Atomic Flows". In: *Logical Methods in Computer Science* 11.1 (2015), 4:1–27. ISSN: 1860-5974. DOI: 10.2168/LMCS-11(1:4)2015. URL: http://arxiv.org/pdf/1502.05860.pdf.

[DS16] Anupam Das and Lutz Straßburger. "On linear rewriting systems for Boolean logic and some applications to proof theory". In: *Logical Methods in Computer Science* 12 (2016).

[DV01] Anatoli Degtyarev and Andrei Voronkov. "Kanger's choices in automated reasoning". In: *Collected papers of Stig Kanger with Essays on his Life and Work*. Springer, 2001, pp. 53–67.

[GAL92] Georges Gonthier, Martín Abadi, and Jean-Jacques Lévy. "The geometry of optimal lambda reduction". In: *Proceedings of the 19th ACM SIGPLAN-SIGACT symposium on Principles of programming languages*. ACM, 1992, pp. 15–26. ISBN: 0897914538.

[Gen64] Gerhard Gentzen. "Investigations into logical deduction". In: *American philosophical quarterly* 1.4 (1964), pp. 288–306. ISSN: 0003-0481.

[GG08] Alessio Guglielmi and Tom Gundersen. "Normalisation Control in Deep Inference Via Atomic Flows". In: *Logical Methods in Com-*

puter Science 4.1 (2008), 9:1–36. ISSN: 1860-5974. DOI: 10.2168/LMCS-4(1:9)2008. URL: http://arxiv.org/pdf/0709.1205.pdf.

[GGP10] Alessio Guglielmi, Tom Gundersen, and Michel Parigot. "A Proof Calculus Which Reduces Syntactic Bureaucracy". In: *Proceedings of the 21st International Conference on Rewriting Techniques and Applications (Rta'10)* 6 (2010), pp. 135–150. ISSN: 1868-8969. DOI: 10.4230/LIPIcs.RTA.2010.135. URL: <GotoISI>://WOS:000392014100011.

[GGS10] Alessio Guglielmi, Tom Gundersen, and Lutz Straßburger. "Breaking Paths in Atomic Flows for Classical Logic". In: *Logic in Computer Science (LICS), 2010 25th Annual IEEE Symposium on*. IEEE, 2010, pp. 284–293. ISBN: 1424475880. DOI: 10.1109/LICS.2010.12. URL: http://www.lix.polytechnique.fr/~lutz/papers/AFII.pdf.

[GHP13] Tom Gundersen, Willem Heijltjes, and Michel Parigot. "Atomic lambda calculus: A typed lambda-calculus with explicit sharing". In: *Proceedings of the 2013 28th Annual ACM/IEEE Symposium on Logic in Computer Science*. IEEE Computer Society, 2013, pp. 311–320. ISBN: 0769550207.

[Gir87] Jean-Yves Girard. "Linear logic". In: *Theoretical computer science* 50.1 (1987), pp. 1–101. ISSN: 0304-3975.

[Gir89] Jean-Yves Girard. "Towards a geometry of interaction". In: *Contemporary Mathematics* 92.69-108 (1989), p. 6.

[GM13] Stéphane Gimenez and Georg Moser. "The Structure of Interaction". In: *CSL*. Vol. 23. 2013, pp. 316–331.

[GTL89] Jean-Yves Girard, Paul Taylor, and Yves Lafont. *Proofs and types*. Vol. 7. Cambridge University Press Cambridge, 1989.

[Gug] Alessio Guglielmi. *Deep Inference*. Web Page. URL: alessio.guglielmi.name/res/cos.

[Gug+10] Alessio Guglielmi et al. *Geometric Normalisation with Atomic Flows*. Web Page. 2010. URL: http://cs.bath.ac.uk/ag/t/GNAF.pdf.

[Gug02] Alessio Guglielmi. *On Lafont's counterexample*. Web Page. 2002. URL: http://cs.bath.ac.uk/ag/p/AG5.pdf.

[Gug04] Alessio Guglielmi. *Formalism A*. Web Page. 2004. URL: http://iccl.tu-dresden.de/guglielm/p/AG11.pdf.

[Gug07] Alessio Guglielmi. "A system of interaction and structure". In: *Acm Transactions on Computational Logic* 8.1 (2007). ISSN: 1529-3785. DOI: Artn110.1145/1182613.1182614. URL: <GotoISI>://WOS:000244402200001.

[Gun09] Tom Erik Gundersen. "A general view of normalisation through atomic flows". Thesis. University of Bath, 2009. URL: https://tel.archives-ouvertes.fr/file/index/docid/509241/filename/thesis.pdf.

[HB74] David Hilbert and Paul Bernays. *Grundlagen der Mathematik II*. Springer, 1974.

[He18] Fanny He. "The Atomic Lambda-Mu Calculus". Thesis. University of Bath, 2018.

[Hei10] Willem Heijltjes. "Classical proof forestry". In: *Annals of Pure and Applied Logic* 161.11 (2010), pp. 1346–1366. ISSN: 0168-0072. DOI: DOI10.1016/j.apal.2010.04.006. URL: <GotoISI>://WOS:000280890300003.

[Her30] Jacques Herbrand. "Recherches sur la théorie de la démonstration". Thesis. University of Paris, 1930.

[Her71] Jacques Herbrand. *Logical writings*. Springer, 1971. ISBN: 9027701768.

[Hil25] David Hilbert. "On the Infinite". In: *From Frege to Gödel: a source book in mathematical logic, 1879-1931*. Harvard University Press, 1925, pp. 369–392. ISBN: 0674324498.

[Hug04] Dominic Hughes. "Deep inference proof theory equals categorical proof theory minus coherence". Unpublished Work. Unpublished. 2004.

[Hug06] Dominic JD Hughes. "Proofs without syntax". In: *Annals of Mathematics* 164.3 (2006). 243zl Times Cited:6 Cited References Count:27, pp. 1065–1076. ISSN: 0003-486X. URL: `<GotoISI>://WOS:000251835700009`.

[Hug18] Dominic JD Hughes. "Unification nets: canonical proof net quantifiers". In: *arXiv preprint arXiv:1802.03224* (2018).

[HW13] Stefan Hetzl and Daniel Weller. "Expansion trees with cut". In: *arXiv preprint arXiv:1308.0428* (2013).

[Jeř08] Emil Jeřábek. "Proof Complexity of the Cut-free Calculus of Structures". In: *Journal of Logic and Computation* 19.2 (2008), pp. 323–339. ISSN: 0955-792X. DOI: `10.1093/logcom/exn054`. URL: `http://users.math.cas.cz/~jerabek/papers/cos.pdf`.

[Kan57] Stig Kanger. "Provability in logic, Stockholm studies in Philosophy, vol. 1". In: *Alynquist & Wiksell* (1957).

[Laf89] Yves Lafont. "Interaction nets". In: *Proceedings of the 17th ACM SIGPLAN-SIGACT symposium on Principles of programming languages*. ACM, 1989, pp. 95–108. ISBN: 0897913434.

[Lam90] John Lamping. "An algorithm for optimal lambda calculus reduction". In: *Proceedings of the 17th ACM SIGPLAN-SIGACT symposium on Principles of programming languages*. ACM, 1990, pp. 16–30. ISBN: 0897913434.

[LS05] François Lamarche and Lutz Straßburger. "Naming proofs in classical propositional logic". In: *Typed Lambda Calculi and Applications*. Springer, 2005, pp. 246–261. ISBN: 3540255931.

[McK06] Richard McKinley. "Categorical models of first-order classical proofs". Thesis. University of Bath, 2006.

[McK10] Richard McKinley. "A sequent calculus demonstration of Herbrand's theorem". In: *arXiv preprint arXiv:1007.3414* (2010).

[McK13] Richard McKinley. "Proof nets for Herbrand's Theorem". In: *ACM Transactions on Computational Logic (TOCL)* 14.1 (2013), pp. 1–31. ISSN: 1529-3785. DOI: `10.1145/2422085.2422090`.

[Mil87] Dale A. Miller. "A compact representation of proofs". In: *Studia Logica* 46.4 (1987), pp. 347–370. ISSN: 1572-8730. DOI: `10.1007/BF00370646`. URL: `https://doi.org/10.1007/BF00370646`.

[Min08] Grigori Mints. "Cut elimination for a simple formulation of epsilon calculus". In: *Annals of Pure and Applied Logic* 152.1-3 (2008). 279tw Times Cited:1 Cited References Count:14, pp. 148–160. ISSN: 0168-0072. DOI: `DOI10.1016/j.apal.2007.11.008`. URL: `<GotoISI>://WOS:000254378600010`.

[MJ56] Edward J McCluskey Jr. "Minimization of Boolean functions". In: *Bell system technical Journal* 35.6 (1956), pp. 1417–1444. ISSN: 0005-8580.

[MZ06] Georg Moser and Richard Zach. "The epsilon calculus and Herbrand complexity". In: *Studia Logica* 82.1 (2006), pp. 133–155. ISSN: 0039-3215.

[Par92] Michel Parigot. "Lambda-mu calculus: an algorithmic interpretation of classical natural deduction". In: *Logic programming and automated reasoning*. Springer, 1992, pp. 190–201. ISBN: 354055727X.

[Pla01] Jan von Plato. "A proof of Gentzen's Hauptsatz without multicut". In: *Archive for Mathematical Logic* 40.1 (2001), pp. 9–18. ISSN: 0933-5846.

[Qui52] Willard V Quine. "The problem of simplifying truth functions". In: *The American mathematical monthly* 59.8 (1952), pp. 521–531. ISSN: 0002-9890.

[Qui55] Willard V Quine. "A way to simplify truth functions". In: *The American Mathematical Monthly* 62.9 (1955), pp. 627–631. ISSN: 0002-9890.

[Ral18] Benjamin Ralph. "A Natural Proof System for Herbrand's Theorem". In: *International Symposium on Logical Foundations of Computer Science*. Springer, 2018, pp. 289–308.

[Ret97] Christian Retoré. "Pomset logic: a non-commutative extension of classical linear logic". In: *Typed Lambda Calculi and Applications*. Springer, 1997, pp. 300–318. ISBN: 3540626883.

[RG15] Benjamin Ralph and Alessio Guglielmi. *Confluent and Natural Cut Elimination in Classical Logic*. Web Page. 2015. URL: people.bath.ac.uk/bdr25/papers/natconpres.

[Smu65] Raymond M Smullyan. "Analytic natural deduction". In: *The Journal of Symbolic Logic* 30.02 (1965), pp. 123–139. ISSN: 1943-5886.

[Smu68] Raymond M Smullyan. "Analytic cut". In: *The Journal of Symbolic Logic* 33.4 (1968). C7746 Times Cited:10 Cited References Count:4, pp. 560–564. ISSN: 0022-4812. DOI: Doi10.2307/2271362. URL: <GotoISI>://WOS:A1968C774600006.

[SS04] Charles Stewart and Phiniki Stouppa. "A Systematic Proof Theory for Several Modal Logics". In: *Advances in modal logic* 5 (2004), pp. 309–333.

[SS17] David R Sherratt and Marco Solieri. "Towards Atomic Graphs". Unpublished Work. 2017. URL: https://people.bath.ac.uk/drs32/files/TowardsAtomicGraphs_2017-06-24.pdf.

[Sta79] Richard Statman. "Lower bounds on Herbrand's theorem". In: *Proceedings of the American Mathematical Society* (1979), pp. 104–107. ISSN: 0002-9939.

[Str03] Lutz Straßburger. "Linear logic and noncommutativity in the calculus of structures". Thesis. Technischen Universität Dresden, 2003.

[Str09] Lutz Straßburger. "Some observations on the proof theory of second order propositional multiplicative linear logic". In: *International Conference on Typed Lambda Calculi and Applications*. Springer, 2009, pp. 309–324. DOI: 10.1007/978-3-642-02273-9_23.

[Str17a] Lutz Straßburger. "Combinatorial flows and their normalisation". In: *LIPIcs-Leibniz International Proceedings in Informatics*. Vol. 84. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017. ISBN: 3959770472.

[Str17b] Lutz Straßburger. *Deep Inference, Expansion Trees, and Proof Graphs for Second Order Propositional Multiplicative Linear Logic*. Report.

Computer Science [cs]/Logic in Computer Science [cs.LO]Reports. Inria Saclay Ile de France, 2017. URL: https://hal.inria.fr/hal-01526831.

[Tiu01] Alwen Tiu. "Properties of a logical system in the calculus of structures". Thesis. Technische Universität Dresden, 2001.

[Tiu05] Alwen Tiu. "A system of interaction and structure II: The need for deep inference". In: *arXiv preprint cs/0512036* (2005).

[Tiu06] Alwen Tiu. "A local system for intuitionistic logic". In: *Logic for Programming, Artificial Intelligence, and Reasoning*. Springer, 2006, pp. 242–256. ISBN: 3540482814.

[Web14] Jack Webb. "On Herbrand's Theorem". Thesis. University of Oxford, 2014.

[Wit61] Ludwig Wittgenstein. *Tractatus Logico-Philosophicus*. Routledge and Kegan Paul, 1961.